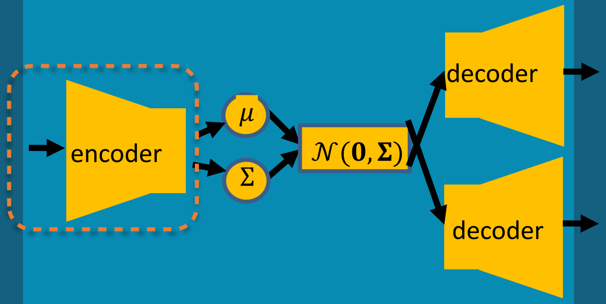
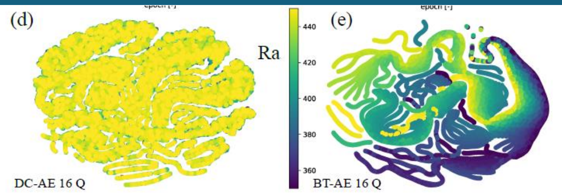
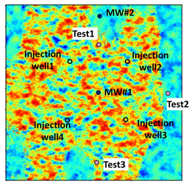
SAND2022-3538C

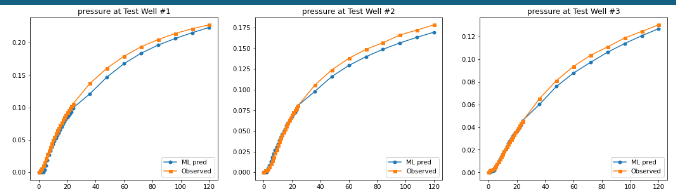
Sandia National Laboratories

Physics-based Deep Learning Driven CO_2 Flow Modeling and Data Assimilation for Real-Time Forecasting









Hongkyu Yoon[1]

Jonghyun Harry Lee[2], Teeratorn Kadeethum[1]

[1]Geomechanics Department, Sandia National Laboratories, NM

[2]Civil and Environmental Engineering, University of Hawaii at Manoa, HI

AAPG CCUS
March 2022

This work was supported by DOE Office of Fossil Energy and Carbon Management project -**Science-informed Machine Learning to Accelerate Real Time (SMART) Decisions in Subsurface Applications-Carbon Storage** & Laboratory Directed Research and Development program at Sandia National Laboratories.

U.S. DEPARTMENT OF ENERGY
NNSA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

Motivation for Deep Learning Based Approach

Two major challenges for high-dimensional forward and inverse problems for real-time forecasting

1. **Computational burdens with matrix calculations (e.g., Jacobian)**

   => Effective dimension reduction

2. **# of forward model simulations for inverse modeling**

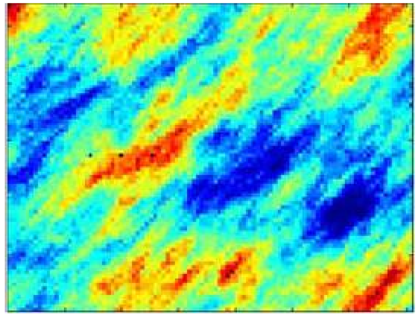
   => ML-driven fast, reduced order predictive modeling

Specific Goals: Machine learning-driven CO_2 modeling by combining fast ML-based forward modeling with (ensemble-based) data assimilation (EnDA), resulting in real-time history matching of CO_2 operations and forecasting CO_2 and pressure plume development

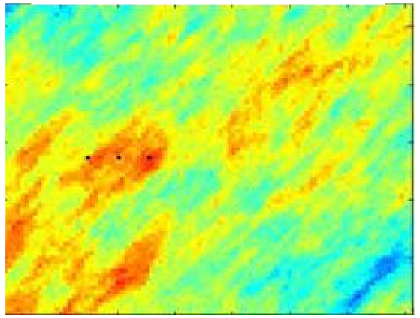
Parameter estimation and uncertainty quantification

History matching
(CO_2 Injection at Cranfield, MS)

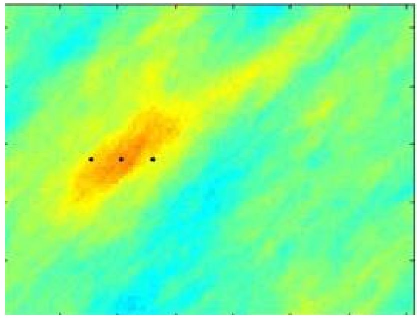
Synthetic Truth



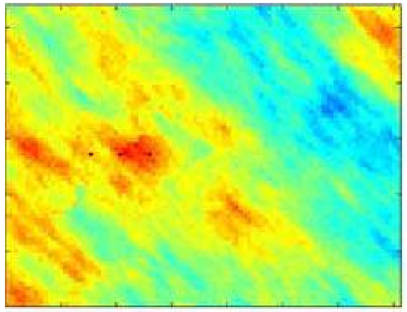
Calibration-constrained NSMC

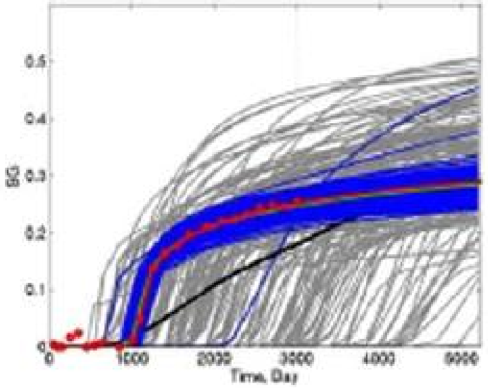


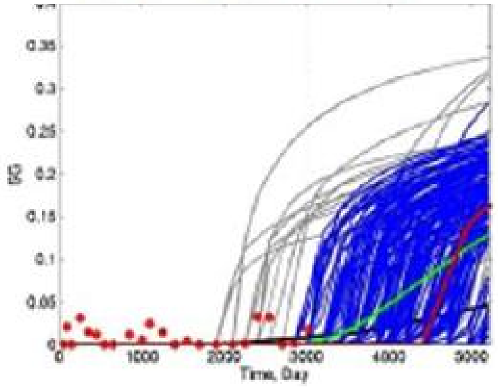
Ensemble-based filtering method



With incorrect prior data







Algorithm
Ensemble Kalman filter
Ensemble smoother
Ensemble smoother with multiple data assimilation
Ensemble Kalman filter with pilot point
ES4 with pilot point
Null-space Monte Carlo[b]
Multiple calibration-constrained NSMC

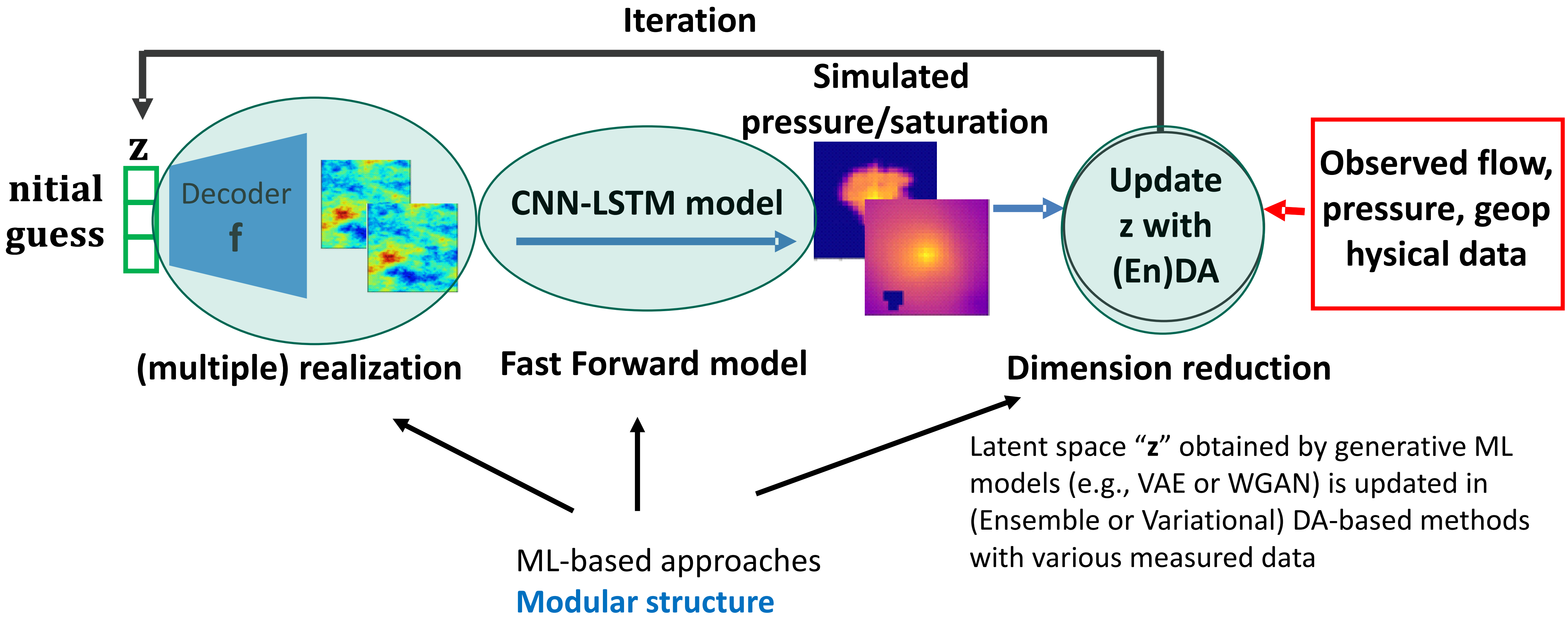
- With limited observation data, solutions with incorrect prior data can match the observed data well ➔ more spatially representative data (e.g., geophysical sensing data, tracer test)
- Another possible solution => more robust ensemble member generation using machine learning

Tavakoli et al. (WRR2013)
Yoon et al. (WRR2013)

Yoon et al. (2007), Yoon & McKenna (2013)
Lee and Yoon et al. (WRR2016)

ML-based Data Assimilation Framework

- Data assimilation in **small nonlinear latent space of unknown parameters with dim(z)**
- Forward model executions can be significantly reduced



Latent space "**z**" obtained by generative ML models (e.g., VAE or WGAN) is updated in (Ensemble or Variational) DA-based methods with various measured data

Kang et al. (in review for WRR), Bao et al. (in prep), Yoon et al. (in prep)

- **ML-based Forward Model**
- ML-based Data Generation
- Data Assimilation
- Summary

ML for Forward Reduced Order Models

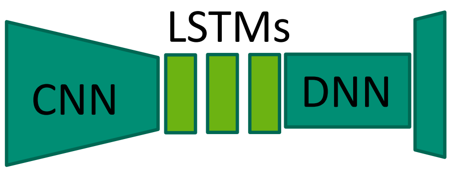
Models for pressure, CO_2 saturation, and water production rate

CNN-LSTM-DNN

Input

- **Permeability and porosity (x,y,z)**
- **Cumulative injection over time**
- **Injection rates/time**
- **Activity binary zone**

ML architecture



CNN, DNN: TimeDistributed

- CNN: Convolutional Neural Network
- LSTM: Long Short Term Memory
- DNN: Dense neural network
- AE: Autoencoder

Output

P & S_{co2} (x,y,z,t)
Q_{prod}(t)

Dimension reduction & interpolation

Physics-Based Loss Functions

- **Loss functions can be constructed through governing equations & physical constraints**
  - We incorporated different terms from governing equations into the loss functions
  - Flux, mass conservation, known quantities are used

Governing equations for two phase flow

$$\frac{\partial(\phi \rho_w S_w)}{\partial t} = \nabla\left(\rho_w \frac{k_{rw}\boxed{\boldsymbol{k}}}{\mu_w}\boxed{(\nabla P_w} - \rho_w g z)\right) + \boldsymbol{q}_w$$

$$\boxed{\frac{\partial(\phi \rho_{nw} S_{nw})}{\partial t}} = \nabla\left(\rho_{nw} \frac{k_{rnw}\boldsymbol{k}}{\mu_{nw}}(\nabla P_{nw} - \rho_{nw} g z)\right) + \boxed{\boldsymbol{q}_{nw}}$$

Loss$= MSE(\hat{P}, P) + MSE(\hat{S}_{nw}, S_{nw}) + MSE(\hat{q}_{pr}, q_{pr})$

$+ \lambda_{flux} * \boxed{MSE(\widehat{Flux}, Flux)}$

$+ \lambda_{mass} * MSE\boxed{\left(\widehat{\frac{\partial(M_{nw})}{\partial t}}, \frac{\partial(M_{nw})}{\partial t}\right)}$

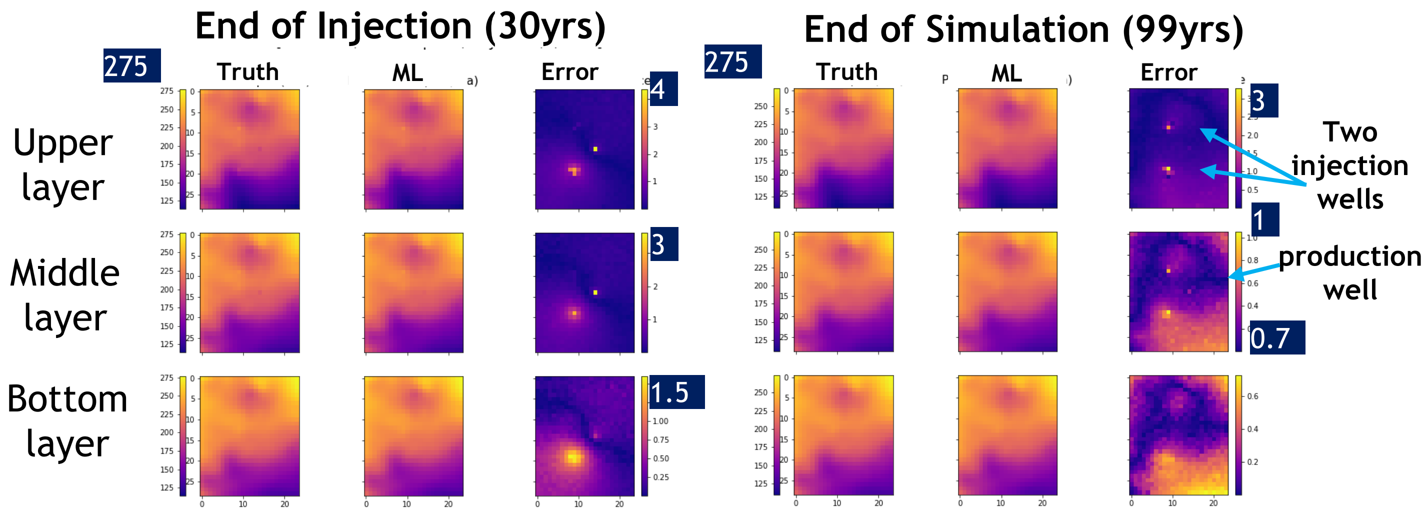
$+ \lambda_{binary} *$ Binary Crossentropy $(\hat{S}_{nw}, S_{nw})$

$+ \lambda_{bhp} * \boxed{MSE(\hat{P}_{bhp}, P_{bhp})} + \lambda_{pr} * \boxed{MSE(\hat{P}_{bhp}, P_{bhp})}$

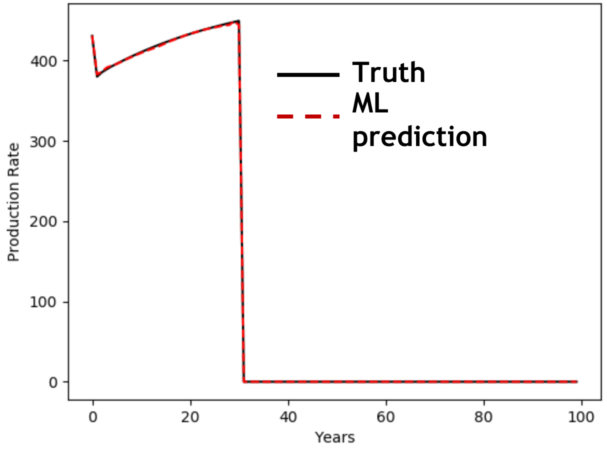
MSE: Mean Square Error

Results – Pressure, CO_2 Saturation & Production Rate

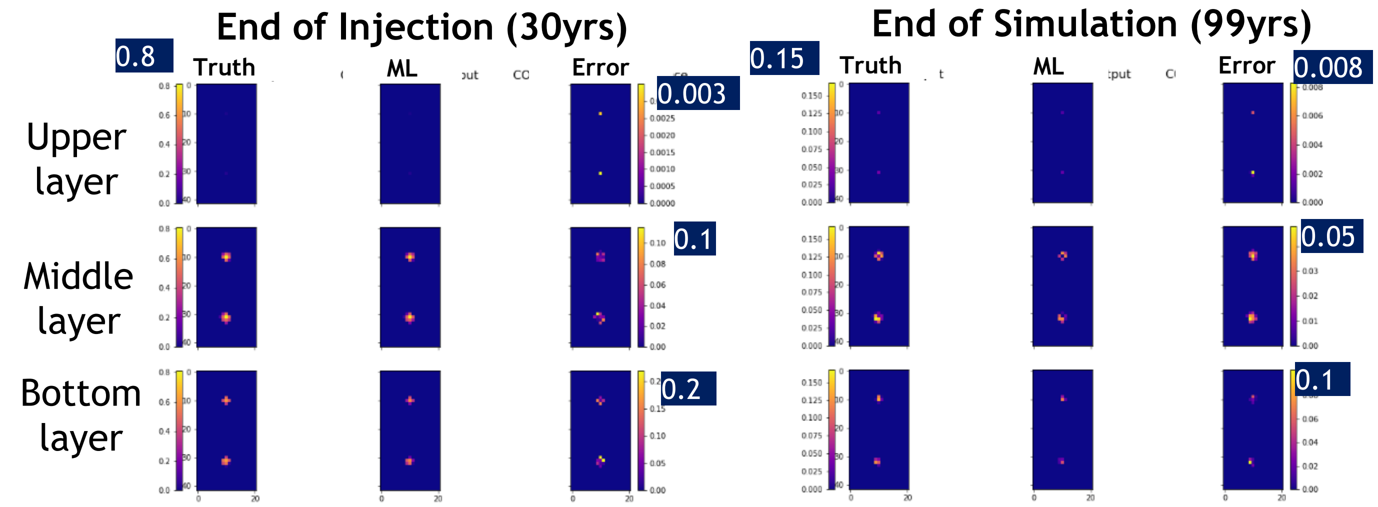
Pressure



Production Rate



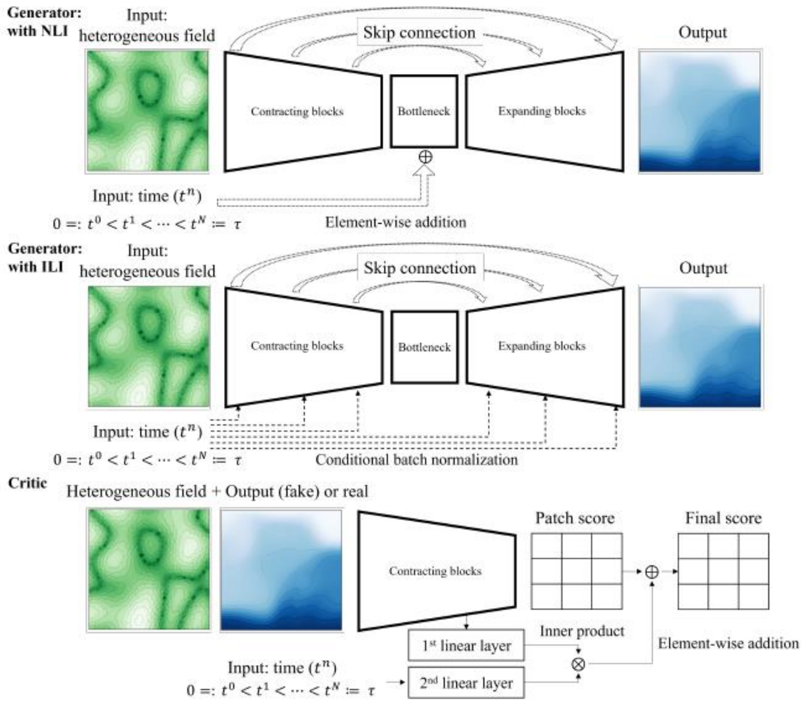
CO_2 Saturation

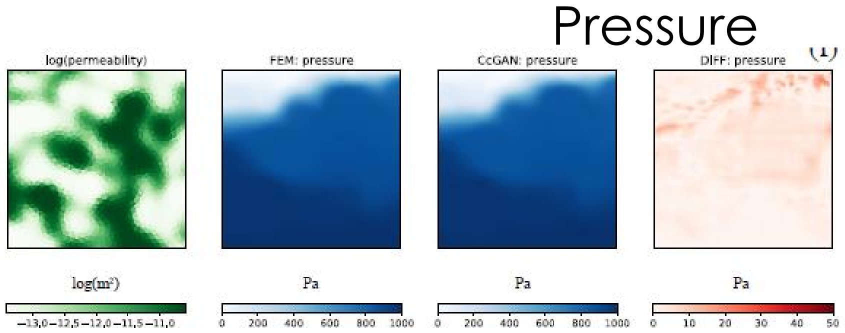


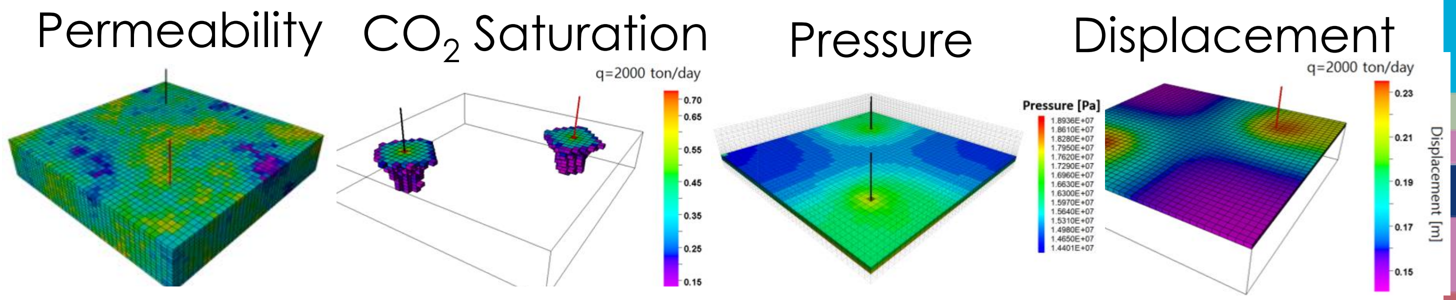
- Trained models has high prediction accuracy for all quantities

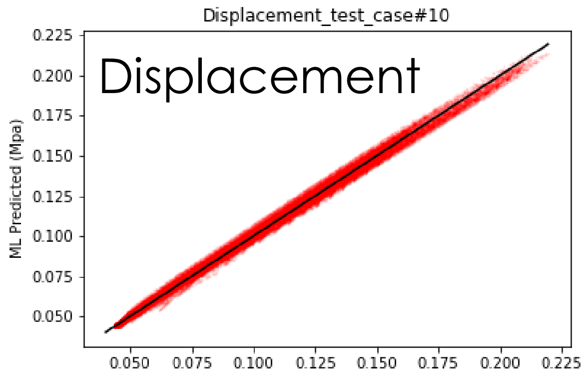
ML approaches for coupled poro-elasticity processes

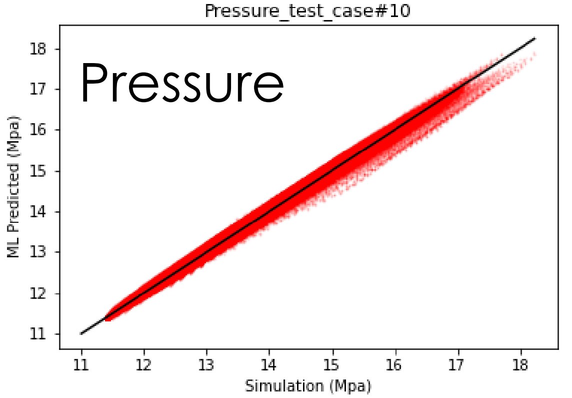
- Continuous conditional generative adversarial networks (CcGAN) for time-dependent PDEs
- CNN-LSTM-DNN reduced order modeling for coupled processes

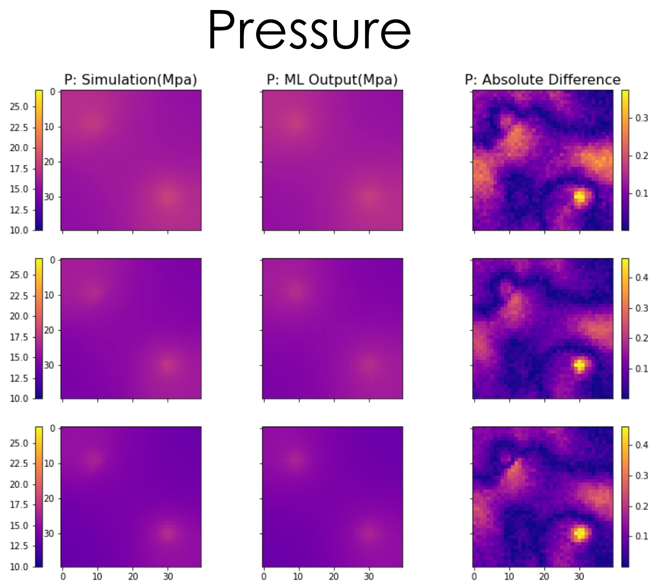










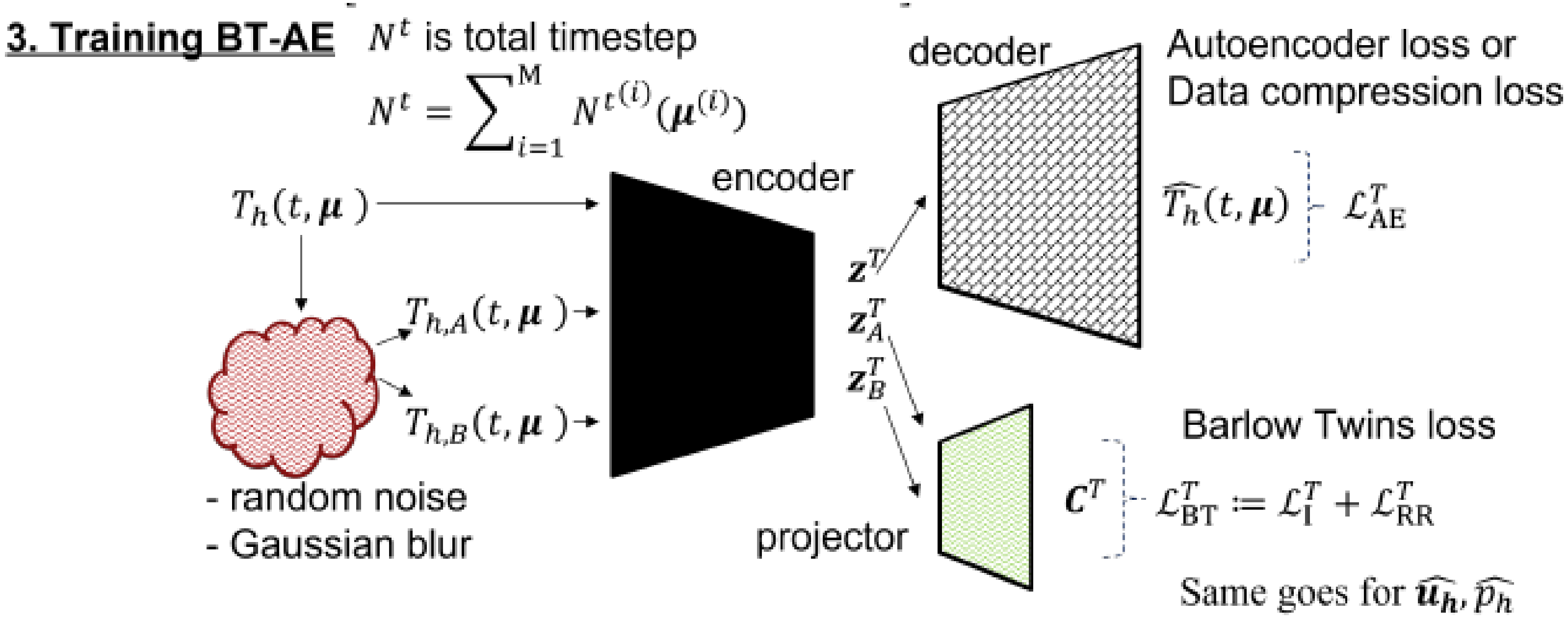


Kadeethum et al. (AWR, 2022; Sci. Rep. 2022)

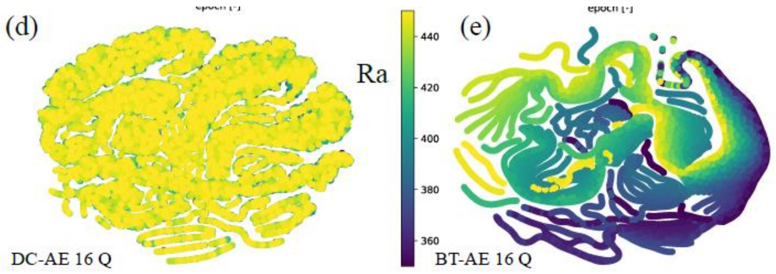
Yoon et al. (in prep for GRL)

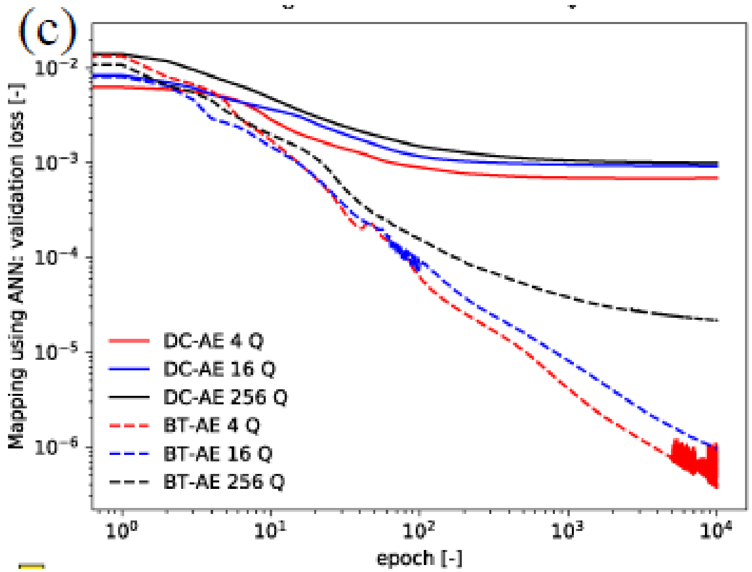
Ongoing Progress

Self supervised ML (Barlow Twins)



DC-AutoEncoder BT-AE



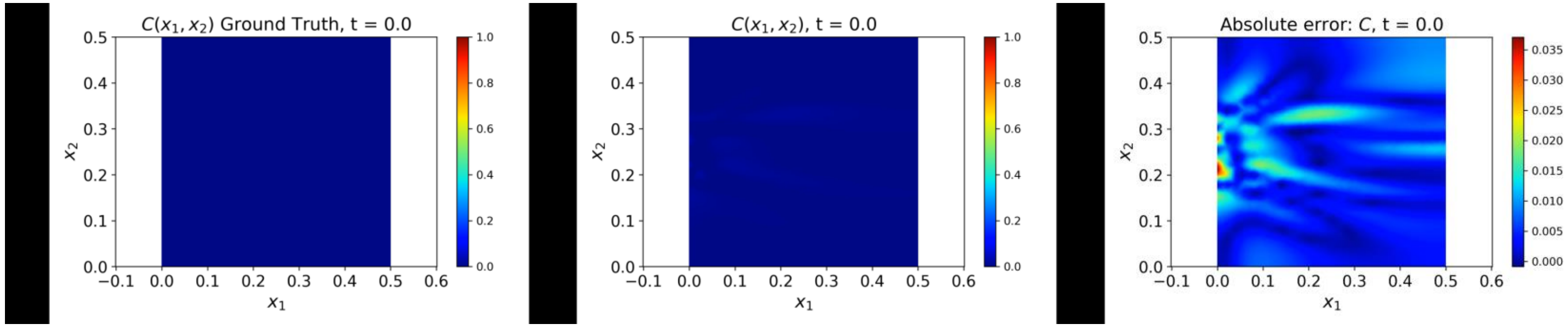


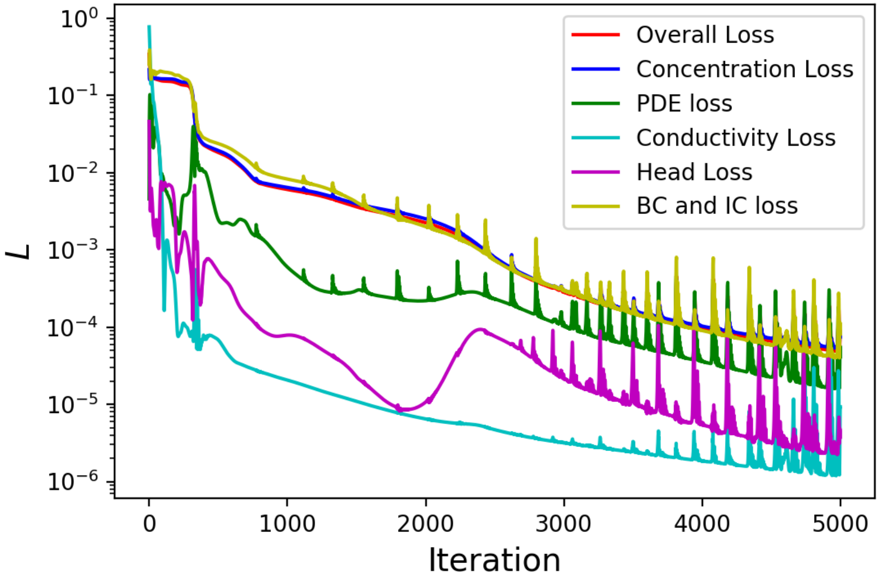
Kadeethum et al. (Sci. Rep., Accepted)

Physics-Informed Neural Networks (PINNs) for PDEs

- A form of neural networks known as **Physics-Informed neural networks** (PINN) to solve **partial differential equations** (PDEs) involved in fluid flow and reactive transport.
- A main idea of PINNs is to **incorporate governing equations of physics** in the form of **partial differential equations** (PDEs) **into the loss** via automatic differentiation (AD)

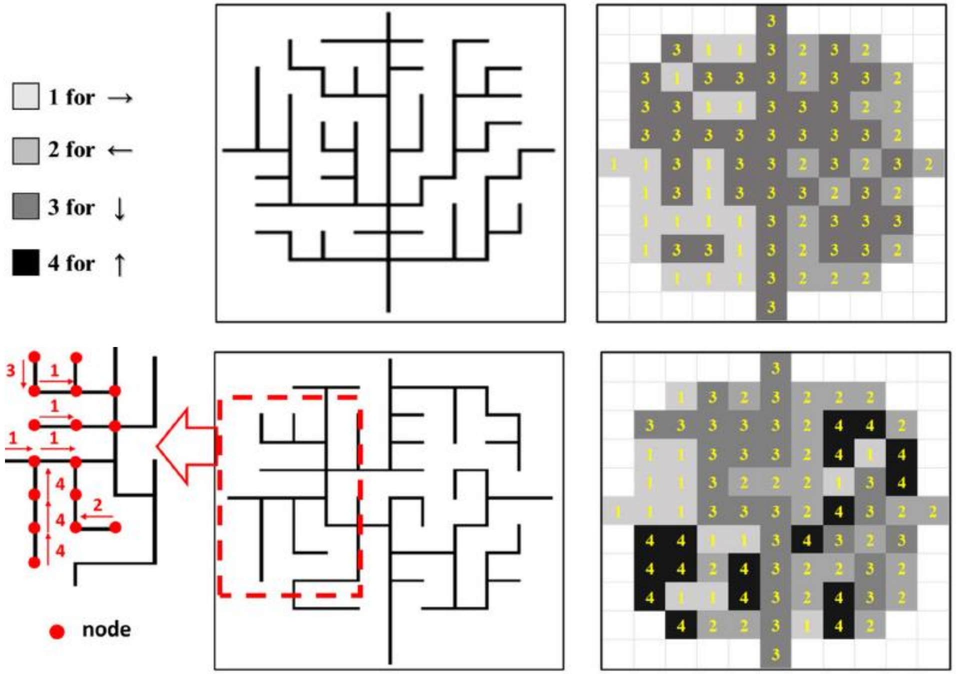
Input: Concentration data + head loss and conductivity +
Advection-Diffusion-Reaction equation + Darcy Equation
Prediction: Permeability field is estimated inversely



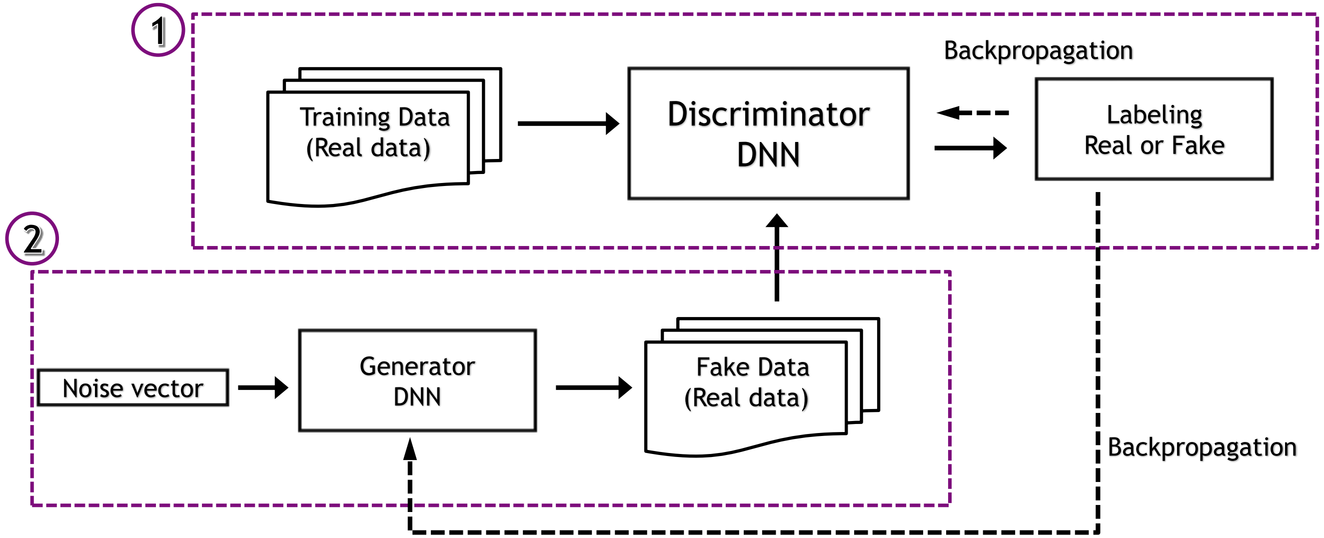


- ML-based Forward Model
- **ML-based Data Generation**
- Data Assimilation
- Summary

Connectivity-Informed Drainage Network Generation



Generative Adversarial Networks (GANs)

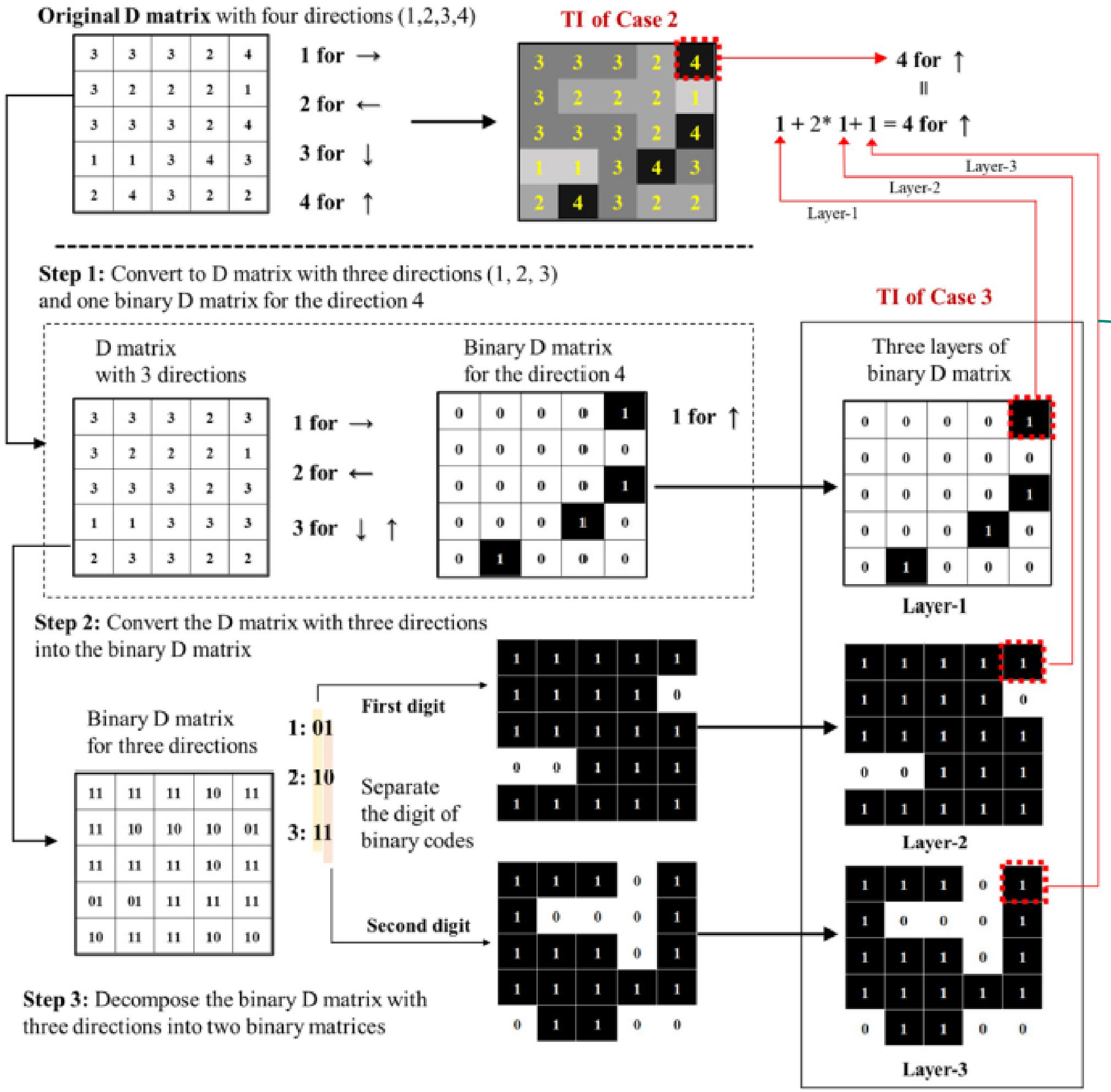


- **Generator (*G*)**: try to fool the discriminator by generating real-looking images from a noise sample
- **Discriminator (*D*)**: try to distinguish between real and fake images

- Training GANs: two player's game (Goodfellow et al., NIPS, 2014)
- Standard GAN is prone to mode collapse & unstable training
- Very active research topics
  - Better loss functions, more stable training (Wasserstein GAN, LSGAN, DCGAN, etc)

Kim et al., Sci. Rep., *2021*

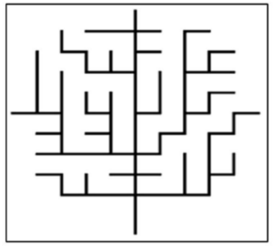
Connectivity-Informed Drainage Networks



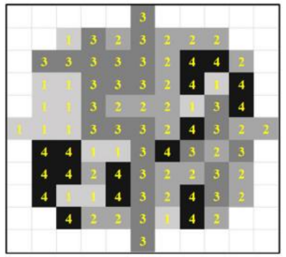
Kim et al., Sci. Rep., *2021*

- **Three cases as training images**

Case 1 – drainage network image



Case 2 – Directional drainage network index image



Case 3 – Three layers of binary D matrix

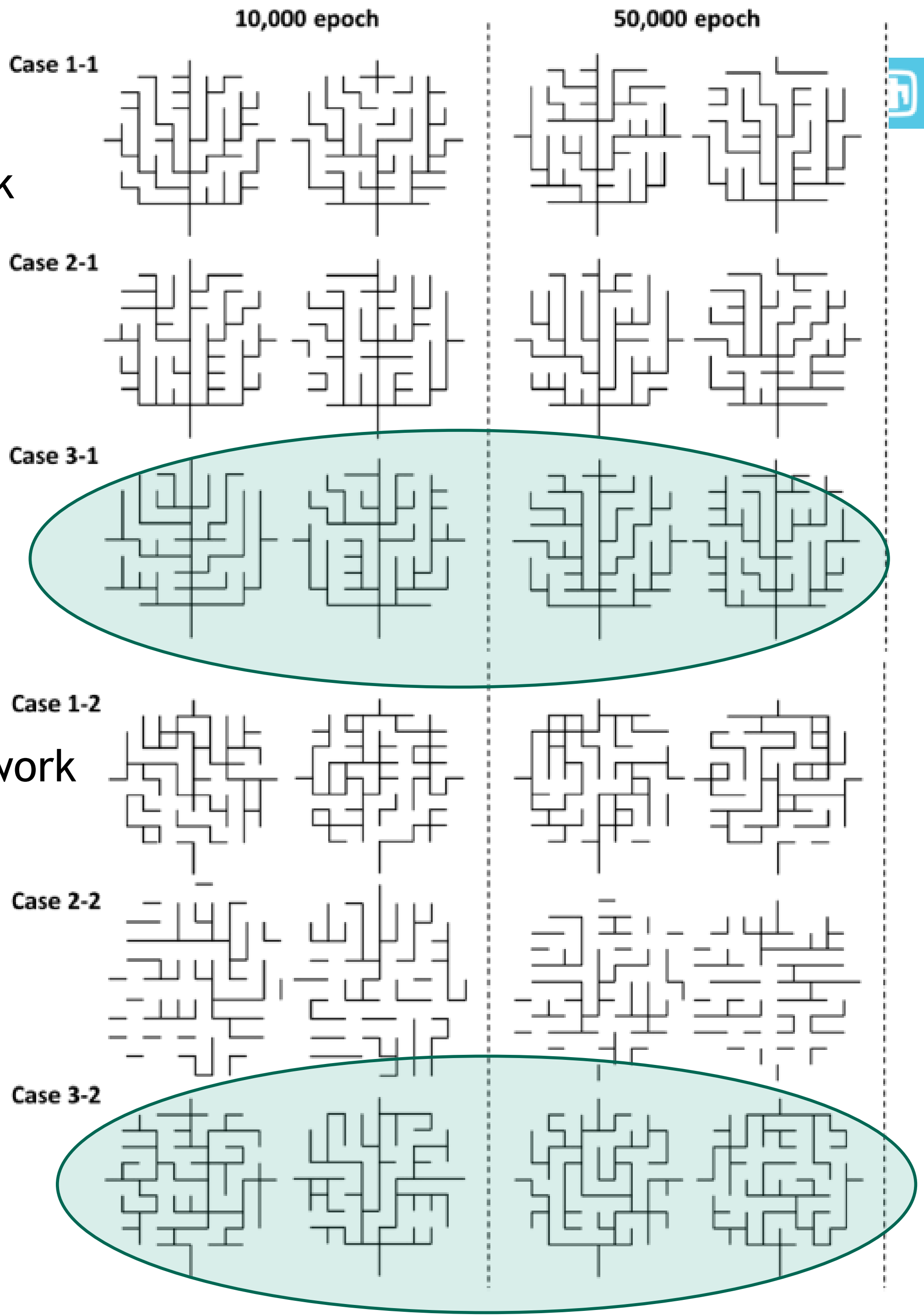
- transform the physical information of the images (i.e., high-frequency features & connectivity between the neighboring nodes) into the efficient binary matrix layers

Generated Drainage Networks

Simple network

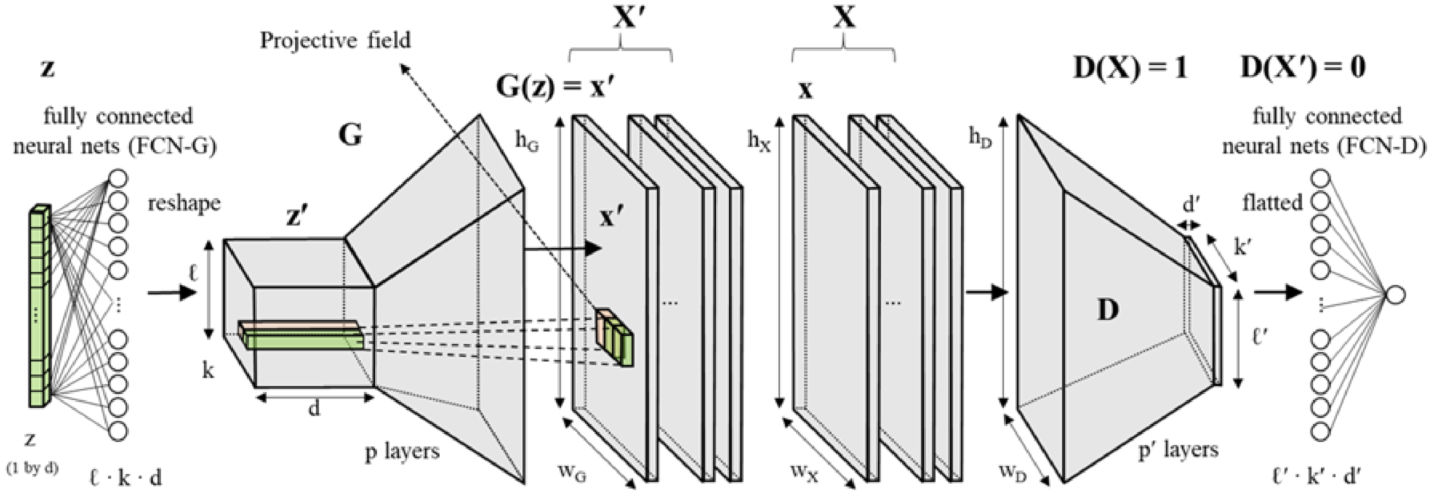
Complex network

- Connectivity-informed binary layers (case 3-1&3-2) outperform other cases
  - Better generation accuracy & computational cost
  - Complex network case demonstrates this more dramatically
  - a type of physics-informed prior knowledge for ML

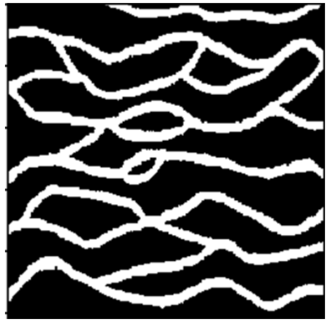


Kim et al., Sci. Rep., *2021*

Spatially Assembled GANs (SAGANs)

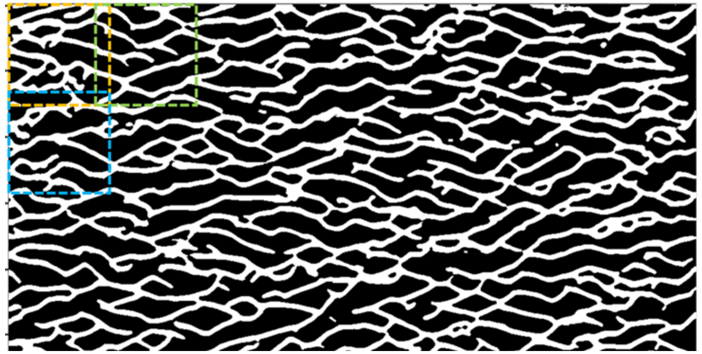


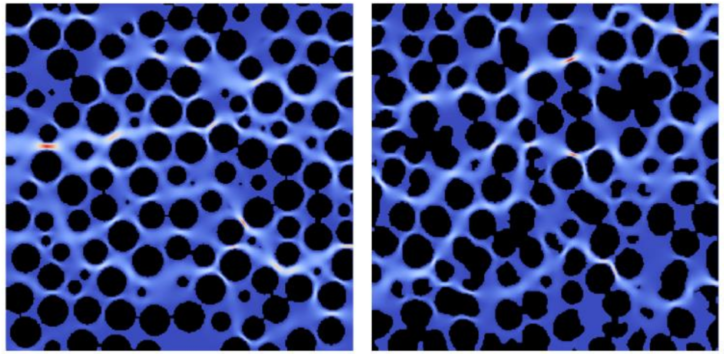
Training image



Training/Generation

Generated image





Velocity field: Pore-scale simulations with training image (TI) and realization

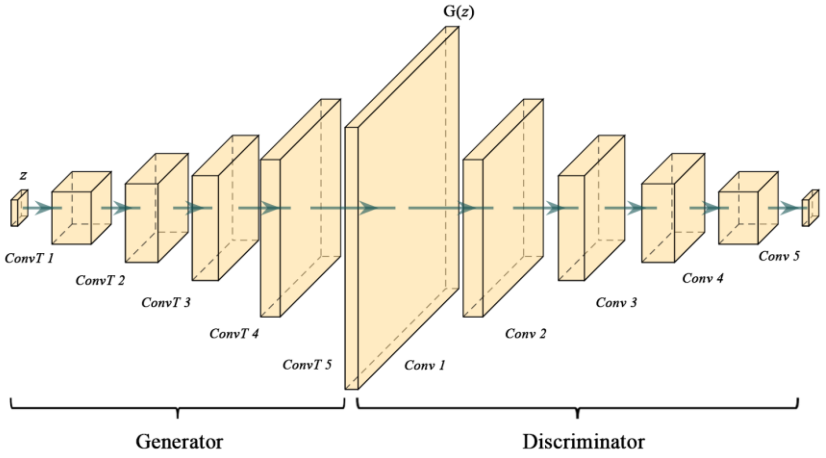
Kim et. al., J. of Contam. Hydro. 2021

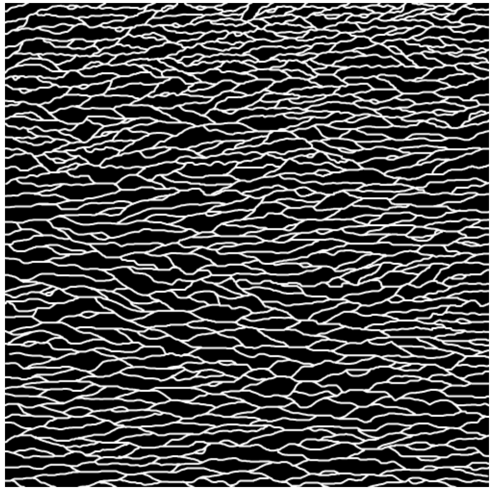
Wasserstein GANs (WGANs)

- As mentioned, "Better loss functions, more stable training"
  - Here we use 1) Wasserstein-1 (so-called Earth-Mover) distance and 2) gradient penalty to ensure Lipschitz (i.e., continuous and differentiable loss function) conditions

$$W(P_r, P_g) = \inf_{\gamma \in \prod(P_r, P_g)} E_{(x,y)\sim\gamma} [\|x - y\|]$$

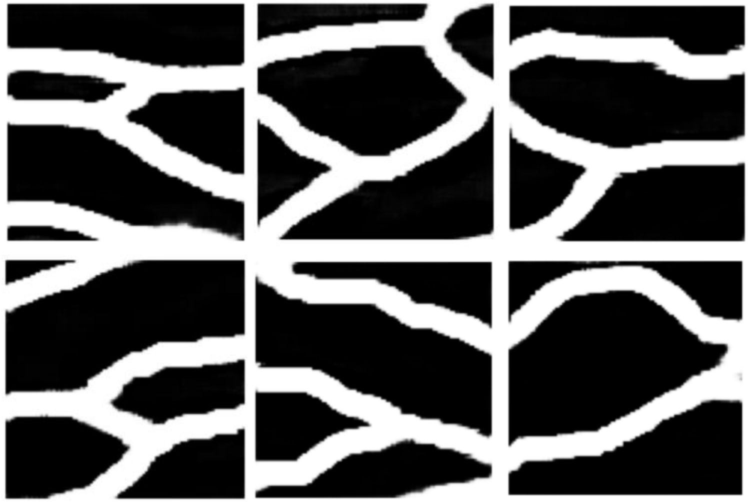
$$L_D = \underbrace{E_{\mathbf{z}\sim P_z}[D(G(\mathbf{z}))] - E_{x\sim P_r}[D(x)]}_{\text{original discriminator loss}} + \underbrace{\lambda E_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]}_{\text{gradient penalty}}$$





Reference training image

Generation



Bao et al., in prep

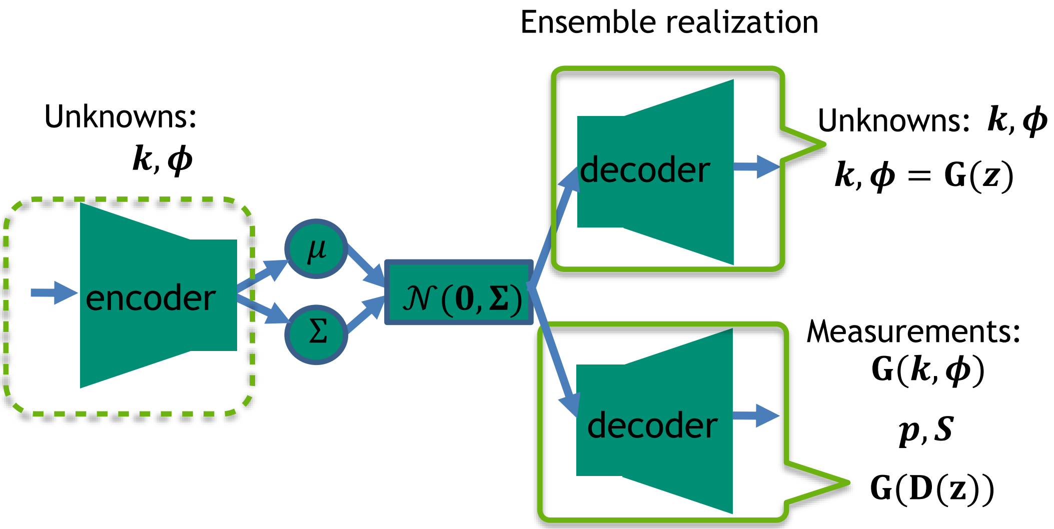
Variational AutoEncoders (VAEs)

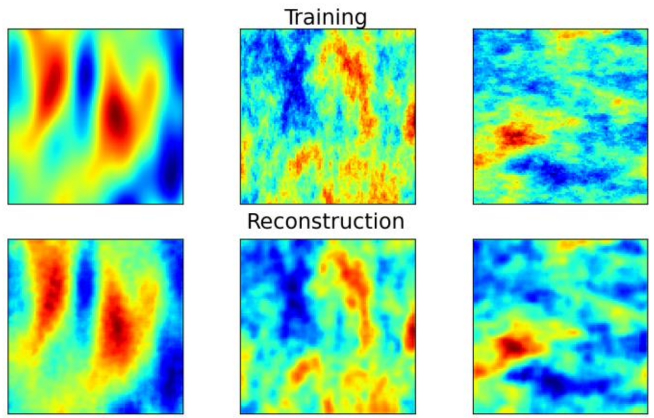
• Nonlinear dimension reduction model:

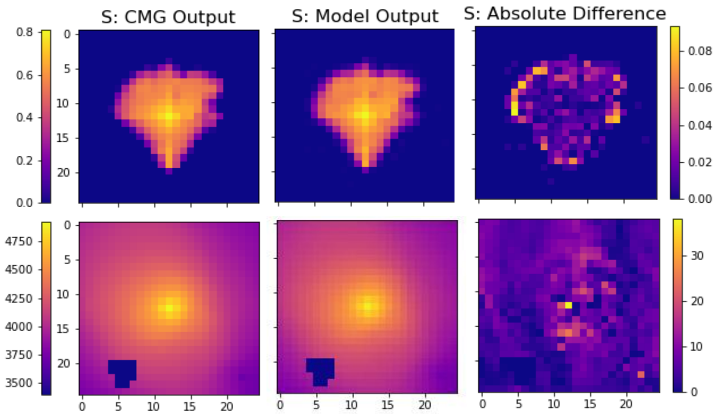
- We have also used VAEs (naïve VAE, β-VAE, VQ-VAE and so on) for data generation.
- VAE can explicitly project data to a smaller space with a simpler (i.e., Gaussian) distribution
- "likelihood" model-based VAEs may be advantageous in some case: relatively easy to 1) train and 2) check the model quality

• Connection to DA

- Data assimilation in **small nonlinear latent space of unknown parameters with dim(z)**
- Only require "**dim(z)**" forward model executions at each iterations instead of dim(m) or dim(obs)

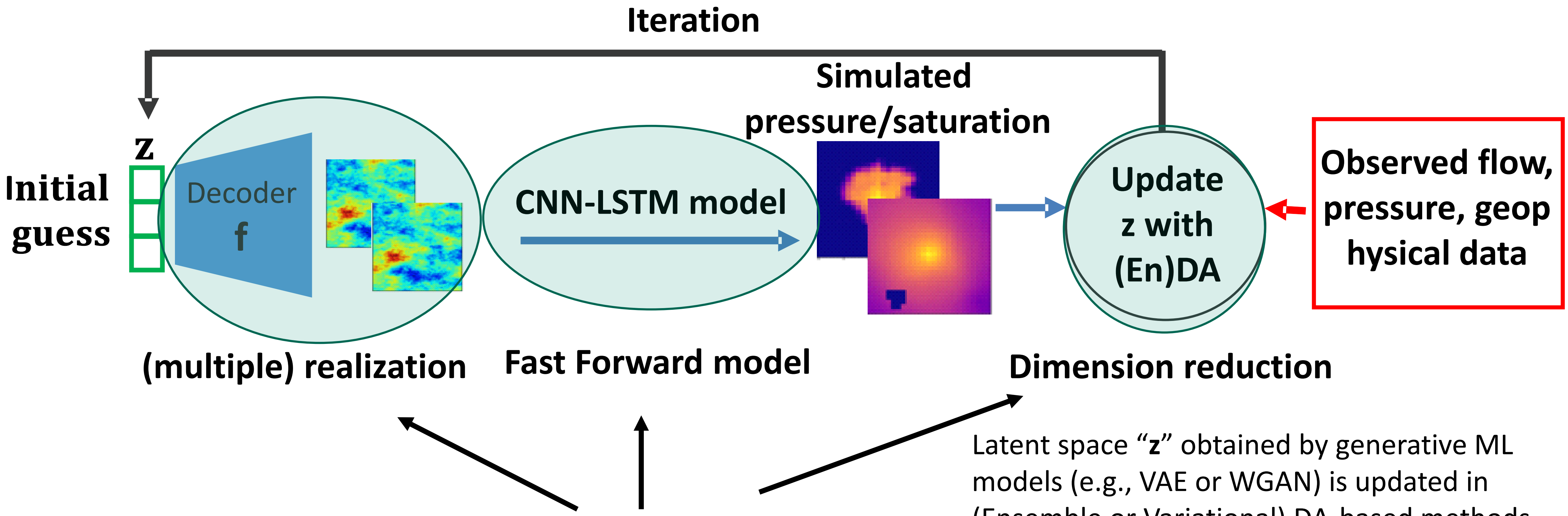






- ML-based Forward Model
- ML-based Data Generation
- **Data Assimilation**
- **Summary**

ML-based Data Assimilation Framework



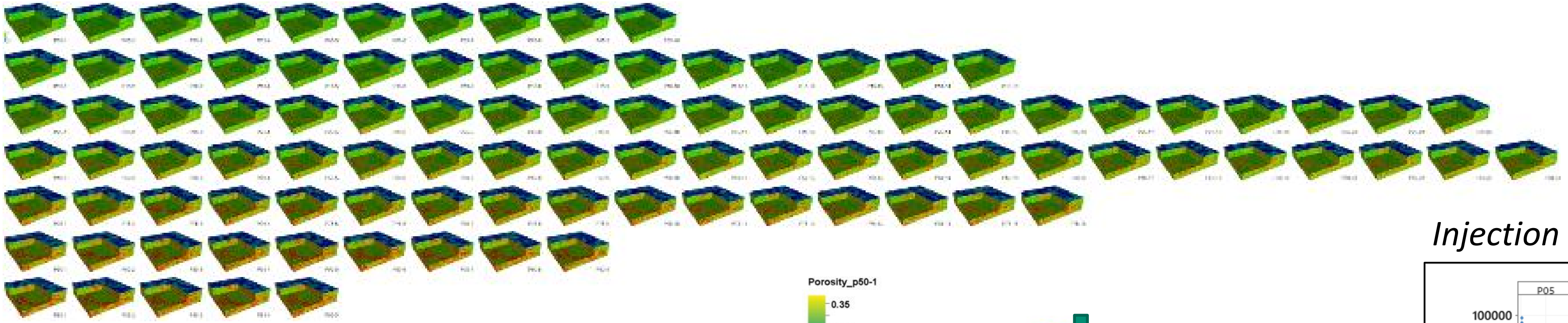
Latent space "**z**" obtained by generative ML models (e.g., VAE or WGAN) is updated in (Ensemble or Variational) DA-based methods with various measured data

ML-based approaches
Modular structure

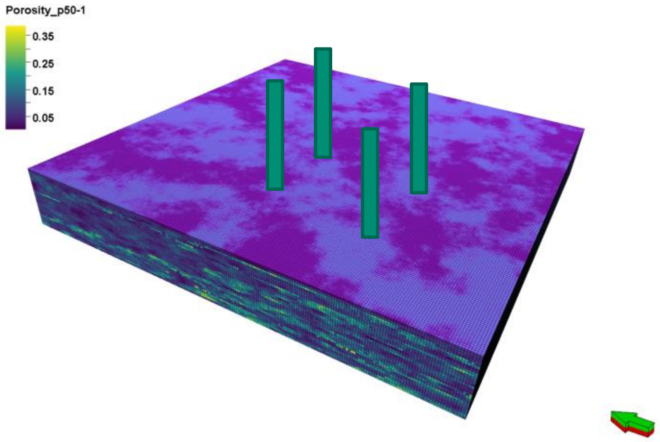
Kang et al. (in review for WRR), Bao et al. (in prep), Yoon et al. (in prep)

Description of the data used

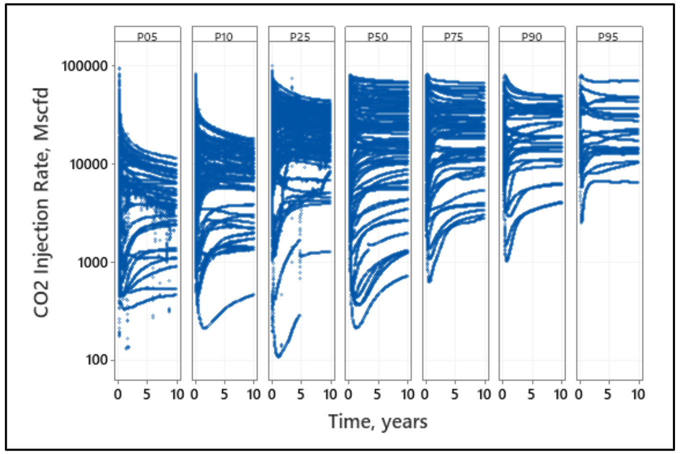
- **High fidelity numerical simulator (CMG) to generate multiphase CO_2 flow in 3D heterogeneous field (DOE SMART-CS project)**
  - Field scale-based permeability & porosity distribution
  - Injection & extraction well operations
  - CO_2 saturation, pressure, and production



Four CO_2 injection wells



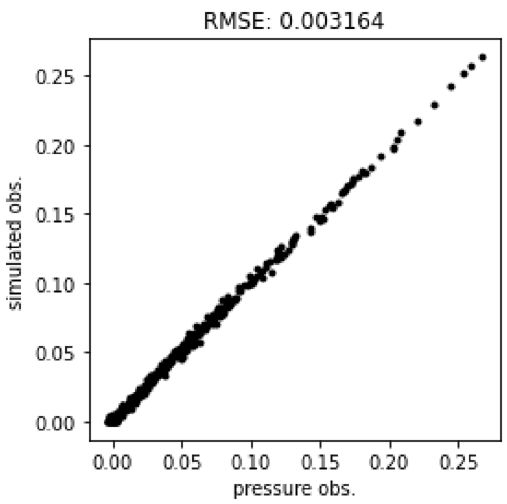
Injection Rate per Well per Realization

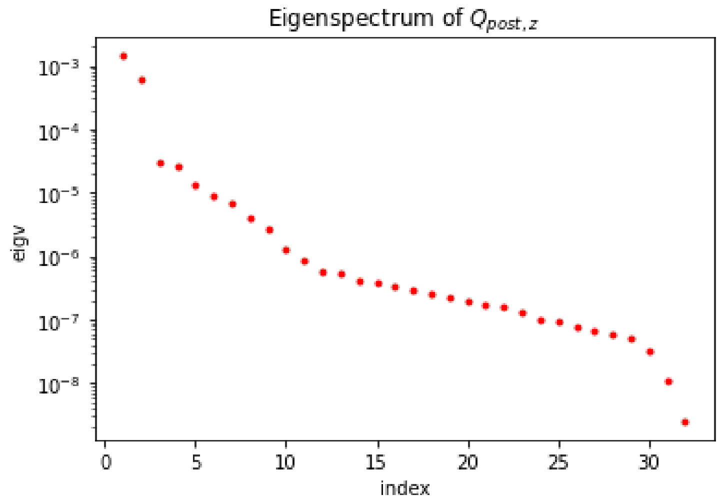


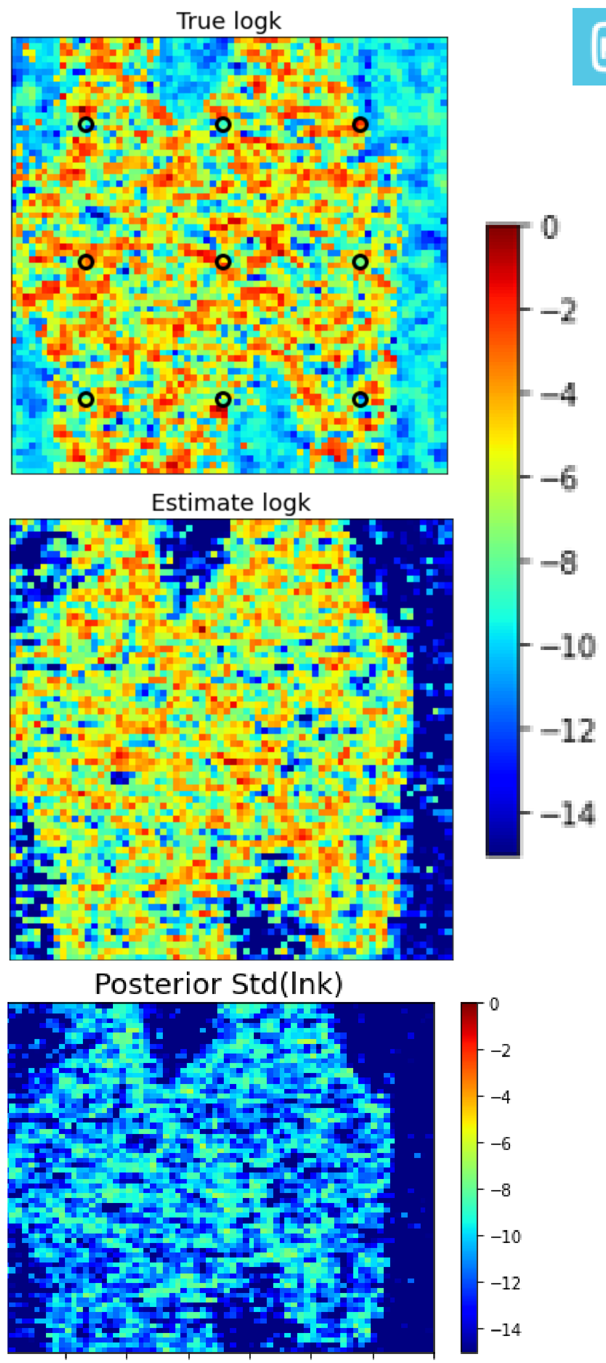
** Visualizations provided by LLNL. Data generı by EERC (Courtesy: Nick Azzolina)*

VAE-Inversion

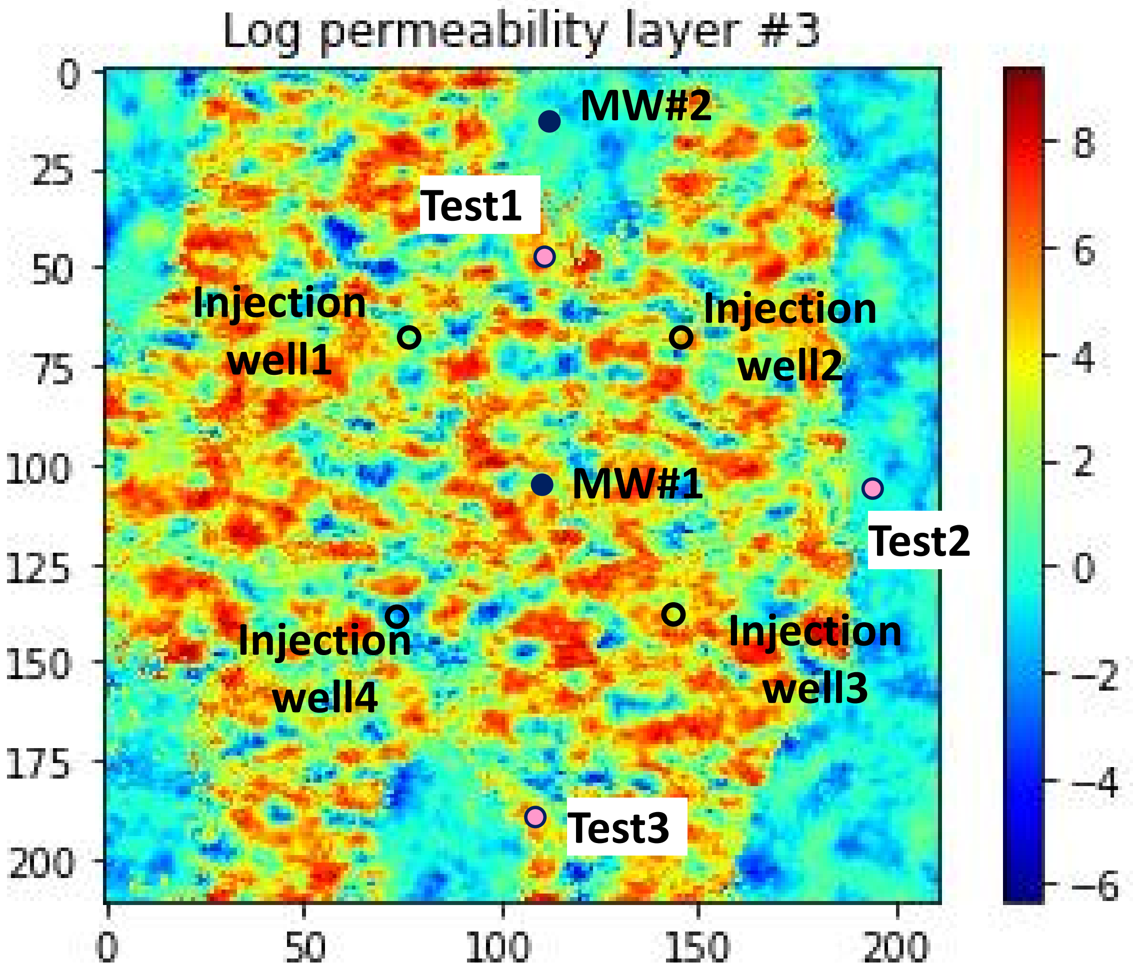
- Use a 2D problem to demonstrate VAE-based inversion
  - the latent space is constructed based on k and Pressure
  - the cost of the trained reduced order model (CNN-LSTM-DNN) ~O(1 sec)
- Inversion example :
  - 2D 71x71 unknown **k** => **z with 32 latent dimension**
  - 9 observation wells for time series pressure & permeability (hard data)
  - Latent space was constructed from training data
  - **Initial guess: Zero mean & STD**
  - Only **~5 min inversion time** on a single core laptop
  - Inversion in the latent space identifies the k structures well!

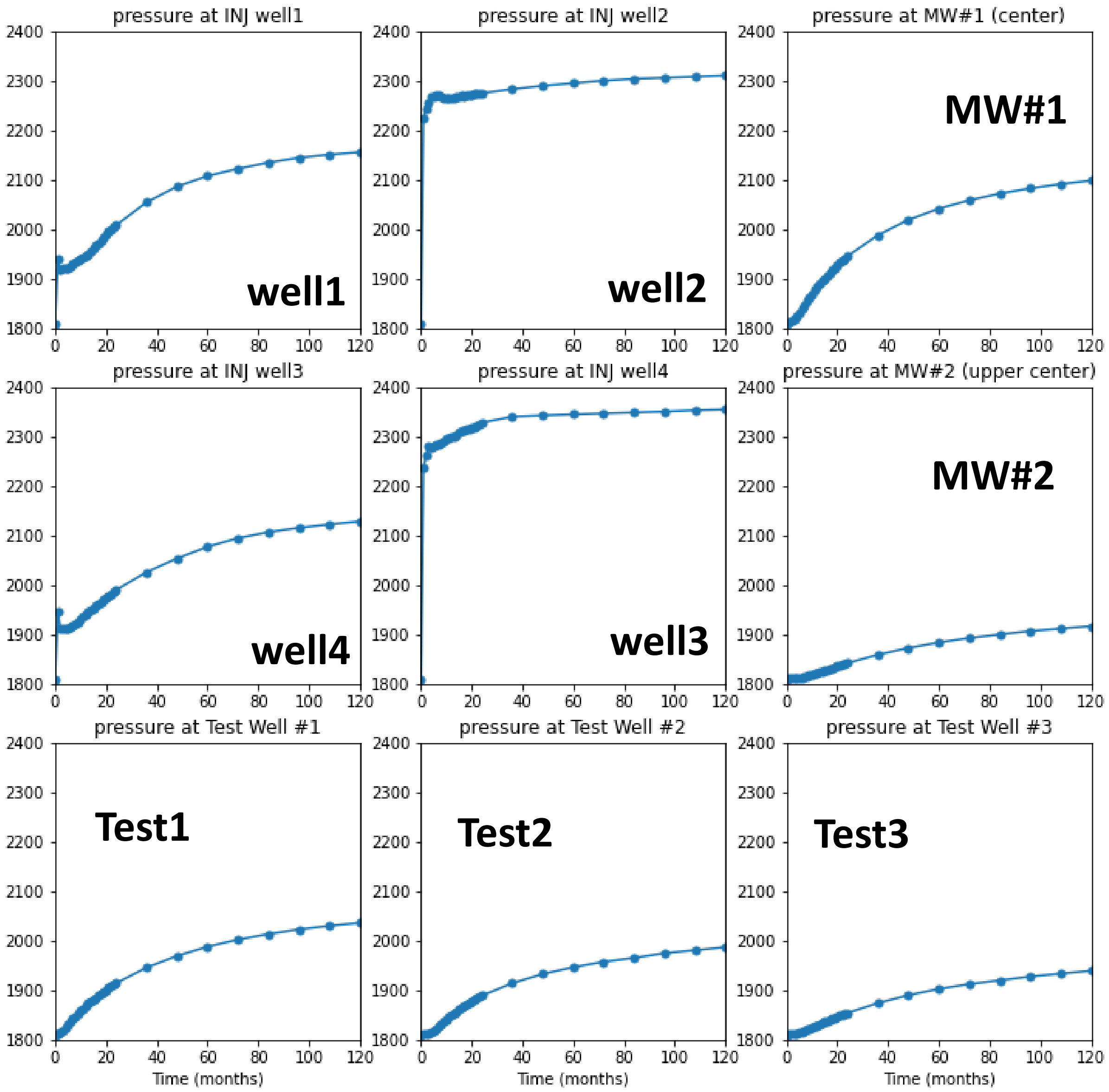






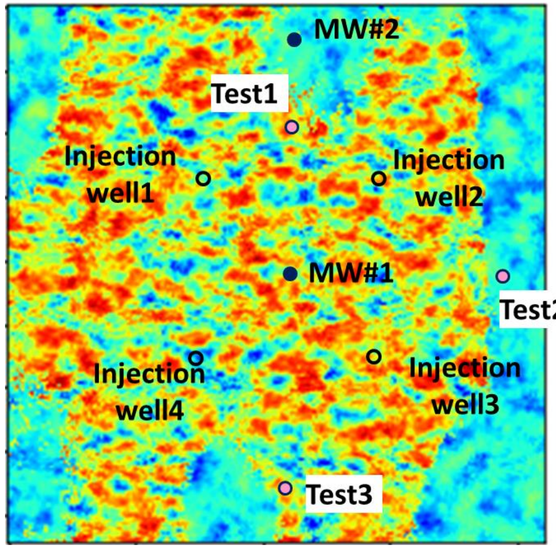
Well locations & pressure profile over time for DA



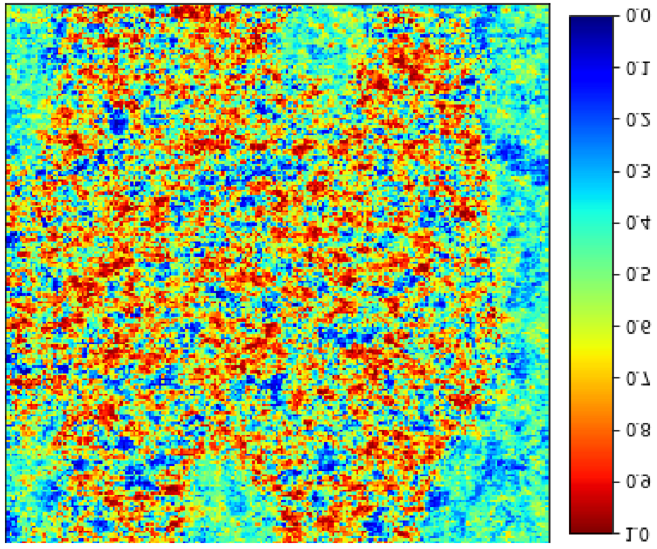


VAE-Inversion

Truth

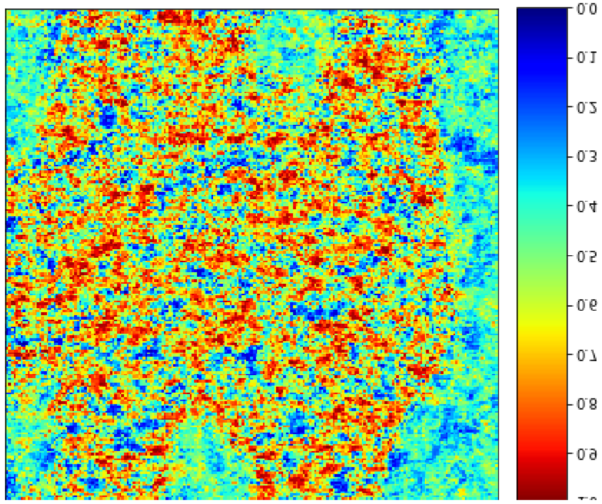


Estimated

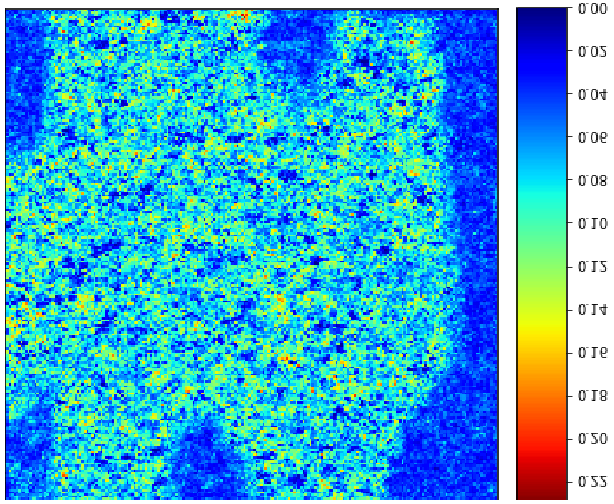


Posterior Analysis (normalized)

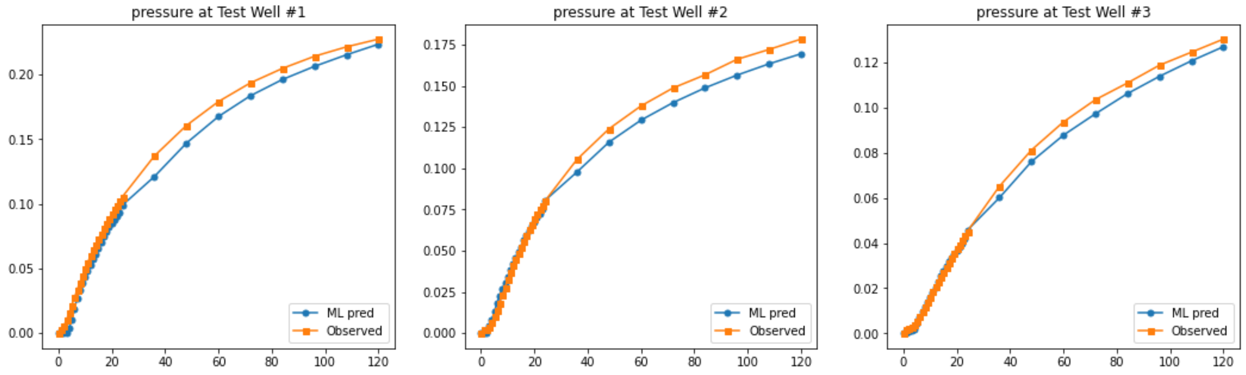
Mean of 1,000 realizations



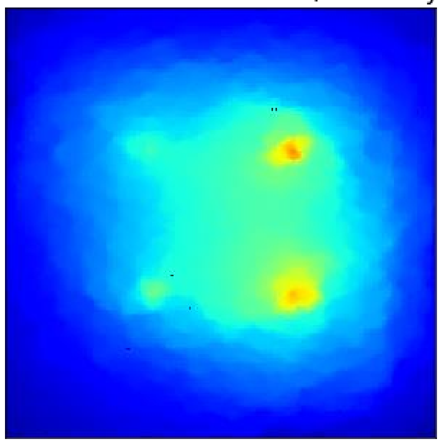
Standard deviation



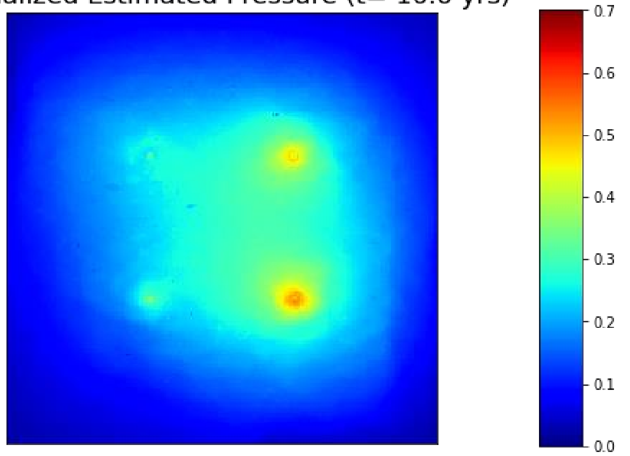
Test #1 / Test #2 / Test #3



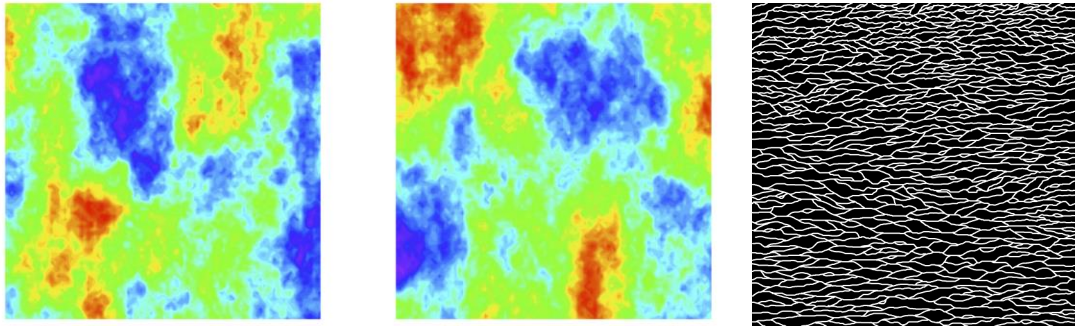
Truth

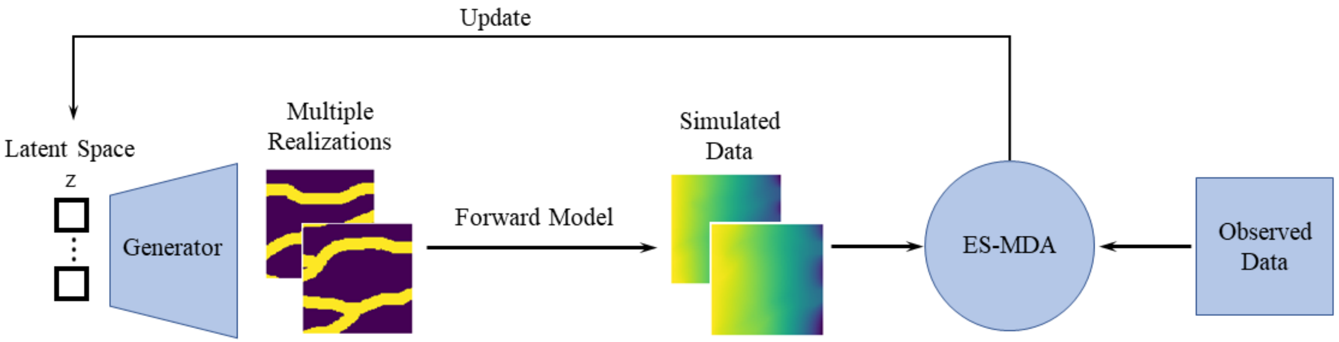


Estimated

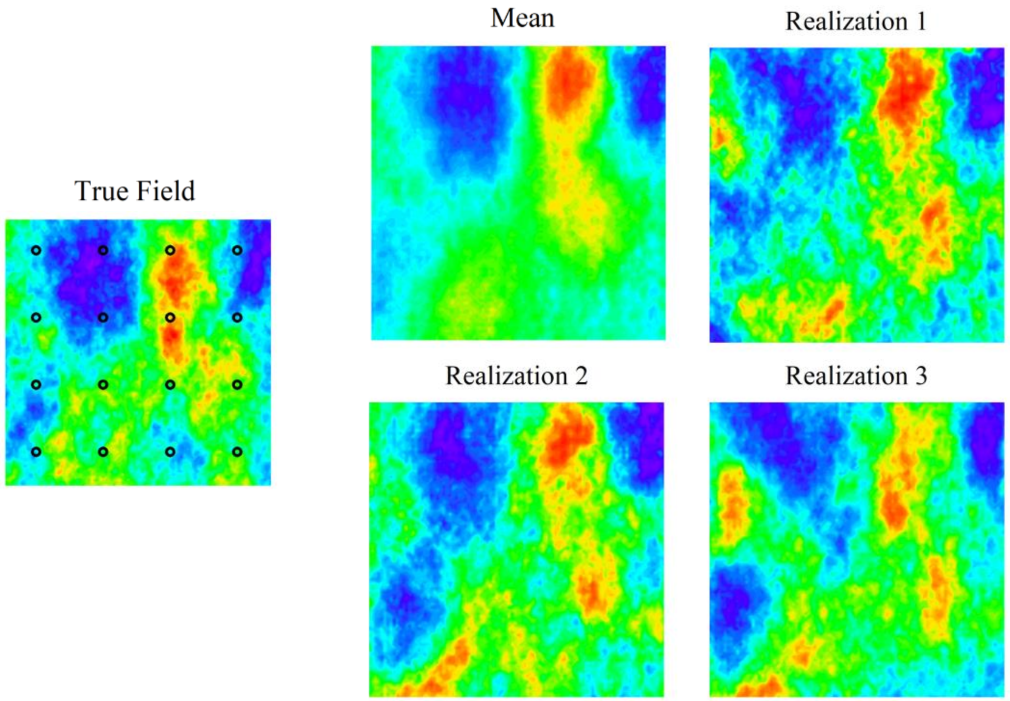


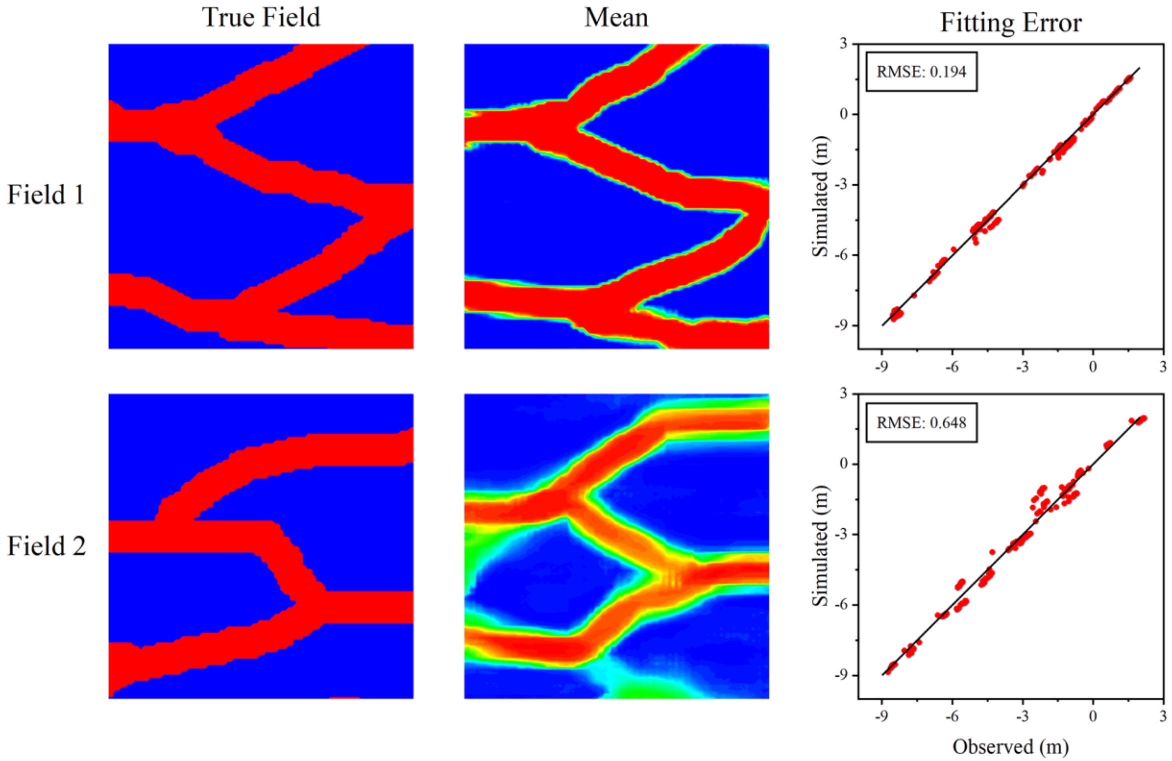
WGAN-based Data Assimilation (ES-MDA)





Training Images for 1) Gaussian and 2) channelized aquifer





Bao et al., in prep

Summary

- Data assimilation in the latent space with deep learning methods (VAE, WGAN) and fast deep learning-based forward modeling can achieve real-time history matching of CO_2 operations and forecasting pressure plume development.

- Latent space optimization including optimal choice of the nonlinear dimension reduction requires further study with more realistic and various types of observed data.

- ML/DL with domain knowledge can lead to dramatic improvement in spatio-temporal data analytics and decision making for optimal monitoring system development.

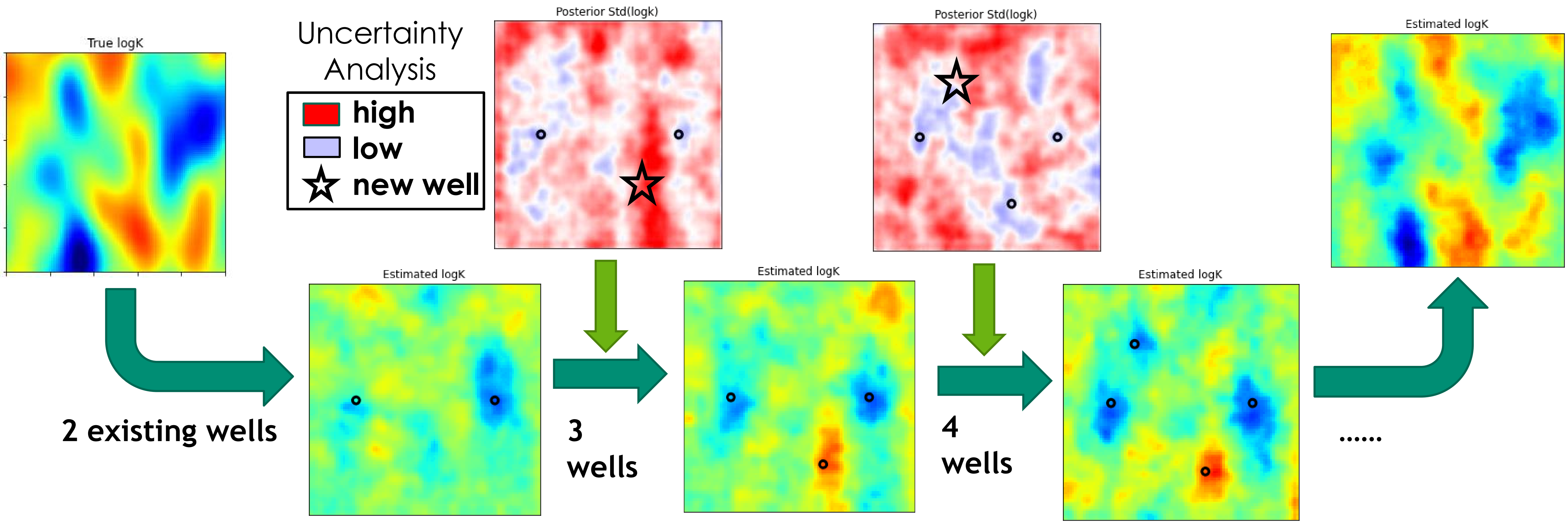
Thank you!

Any questions?

Hongkyu Yoon

hyoon@sandia.gov

Preliminary Result: Optimal Monitoring Well Placement



- By computing posterior covariance and <u>maximize the information gain</u> (e.g, D optimality) in the small latent space, our data assimilation method can accelerate Optimal Experiment Design (OED) problems and identify next "best" well locations