SAND2020-11812PE

SMART-CS Initiative

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

Realtime forecasting of CO_2 flow using variational autoencoder with ensemble-based data assimilation

Task 4

Teams: Hongkyu Yoon (Sandia Lab) & Jonghyun Harry Lee (U of Hawaii at Mānoa)

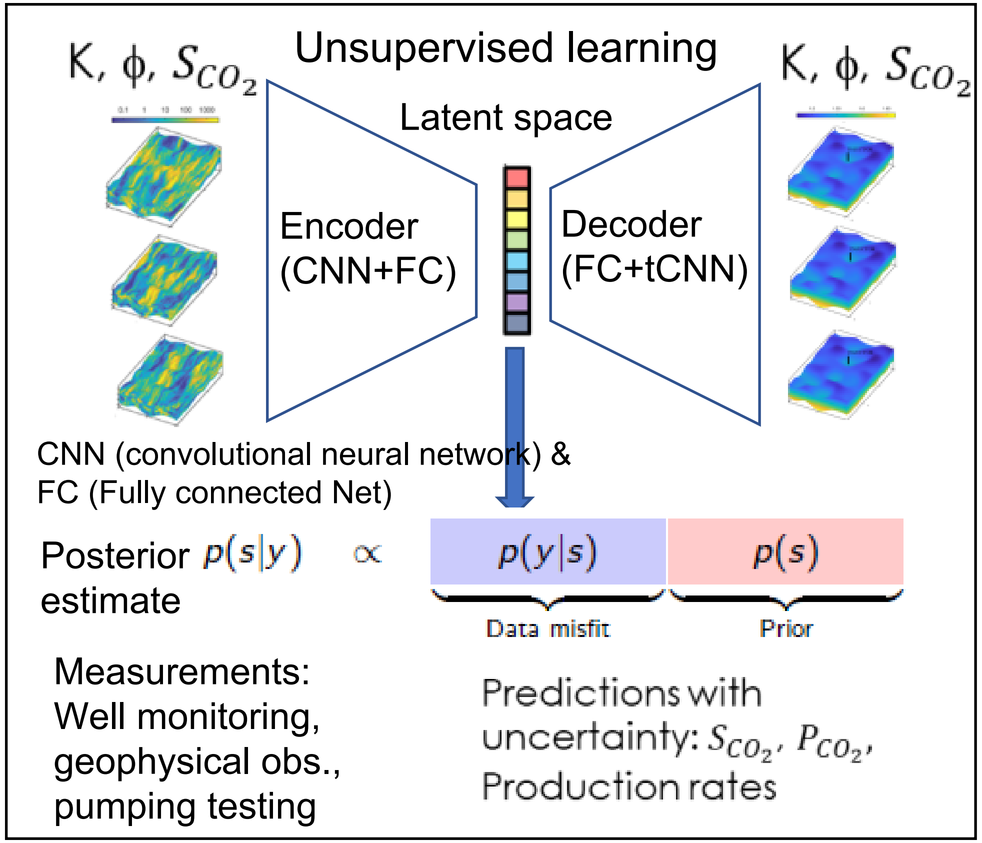
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

U.S. DEPARTMENT OF ENERGY

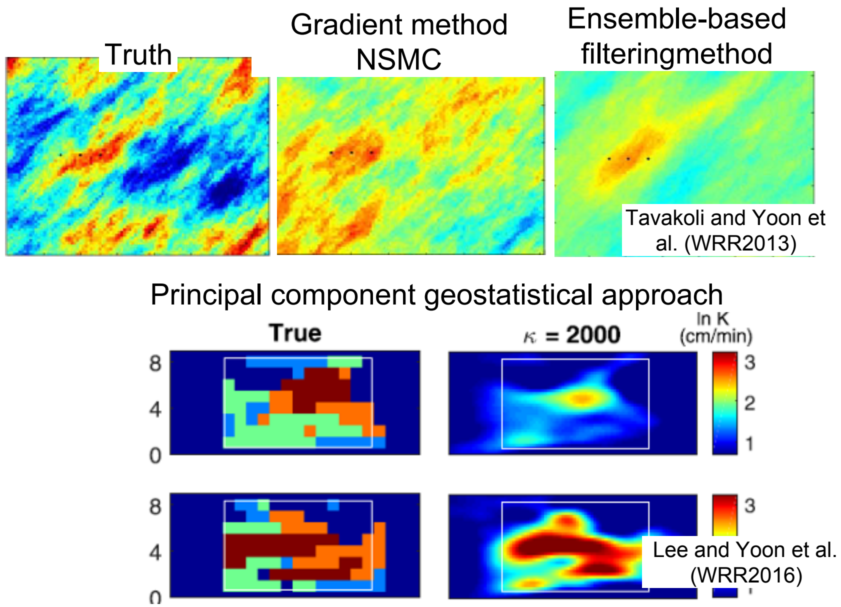
NATIONAL ENERGY TECHNOLOGY LABORATORY

Motivation for Deep Learning Based Approach

Machine learning-driven CO_2 modeling by combining a variational autoencoder (VAE) with ensemble-based data assimilation (EnDA), resulting in real-time history matching of CO_2 operations and forecasting CO_2 and pressure plume development



History matching/Data Assimilation



Challenges: high-dimension problems
- Computational burdens with matrix calculations & # of forward model runs

Variational AutoEncoder-based Inversion

Deep Learning-based nonlinear projection approach to accelerate the stochastic inversion. We use VAE and its decoder to map the permeability k to the latent vector z whose dimension is much smaller than the original dimension of k while ensuring a good approximation accuracy.

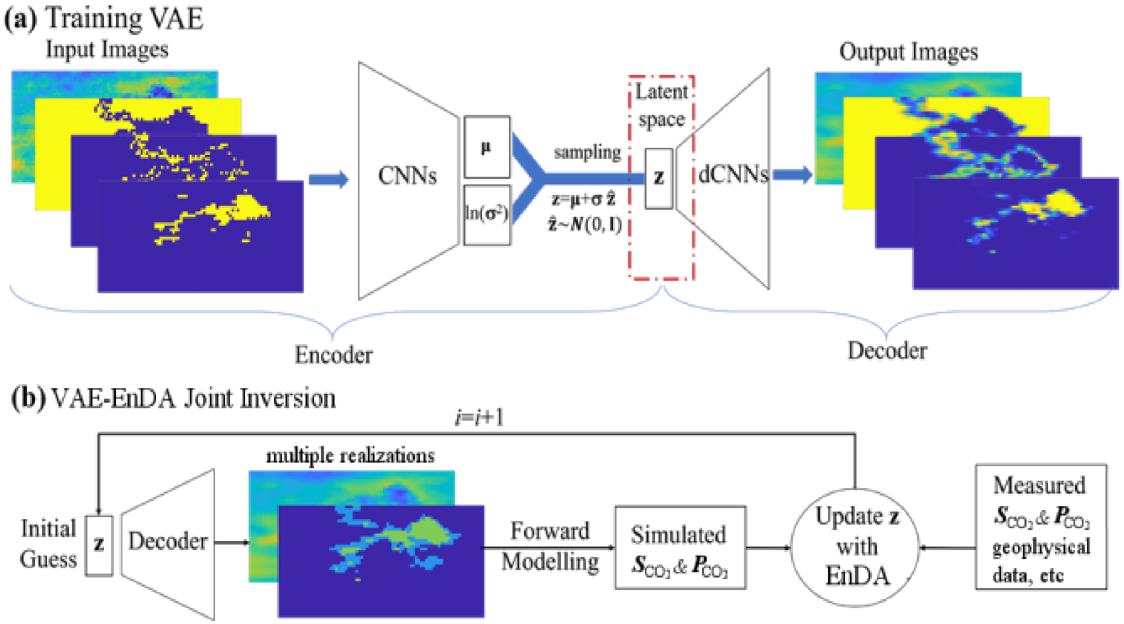
Forward problem: $\mathrm{y} = \mathbf{G}(\mathrm{m})$ with l Gauss Newton iterations at convergence from $\mathrm{m}^0 = \mathrm{m}_{prior}$

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

With any (nonlinear) dimension reduction $\mathbf{G}$ of m

$$y = \mathbf{G}(\mathbf{D}(\mathrm{z})),\ \dim(z) \ll \dim(m)$$

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

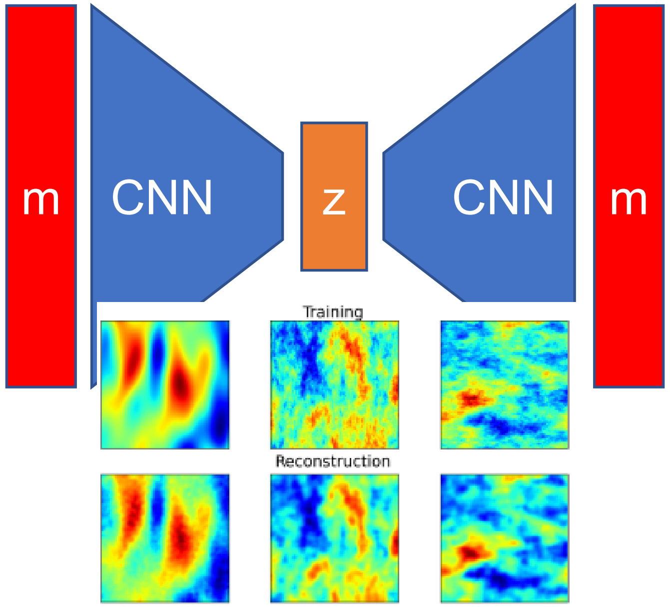


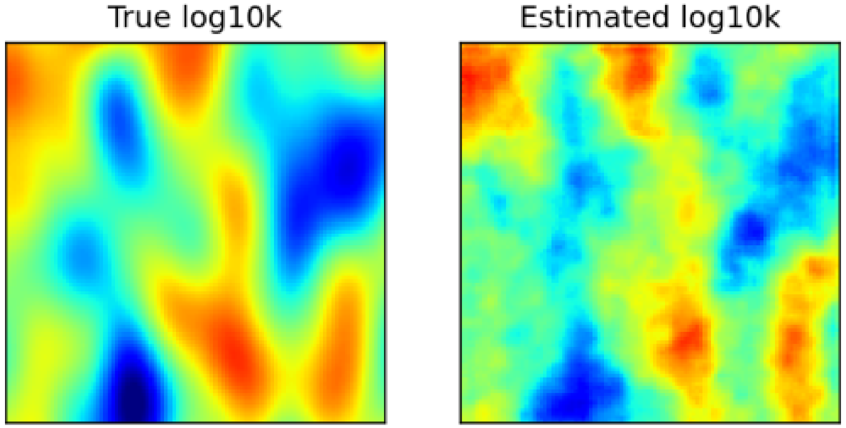
(a) Variational autoencoder constructs a generative model (decoder) for probable permeability fields and CO_2 saturation. (b) Latent space "**z**" obtained by VAE, i.e., deep learning-based encoder will be updated in EnDA-based methods for data assimilation with various measured data.

Preliminary Work

- Use VAE to construct data-driven nonlinear dimension reduction model
  - Only require “dim(z)” forward model executions at each iterations instead of dim(m) or dim(obs)
  - Can encode prior beyond Gaussian

- A nonlinear tomography problem:
  - 10,000 (100x100) unknown **k** => **z** with 32 latent dimension
  - 30 noisy observations
  - **33** forward model runs/iteration to construct Jacobian





Ongoing Progress

Sandia National Laboratories

- Inversion with multi-phase flow modeling examples (2D case first)
- Extension to Data Assimilation with Task 4 problems
- Reduced Order Model (ROM) for particle filtering-type DA - Latent-space-based ROM
- Physics-informed NN approximation to multiphase flow models
  - Physics-based loss function with various DL architecture
  - Physics-informed NN construction for large scale reservoir models