Advanced Machine Learning and Computational Methods

SMART (Science-informed Machine Learning for Accelerating Real-Time Decisions in Subsurface Applications) Phase 2

Jared Schuetter
Senior Data Scientist, Battelle

Alex Tartakovsky
Professor, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign

Chung Shih
Technical Fellow, Senior Strategic Data Scientist, NETL Support Contractor

U.S. DEPARTMENT OF ENERGY

NATIONAL ENERGY TECHNOLOGY LABORATORY

Disclaimer

This project was funded by the United States Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Authors and Contact Information

***Jared Schuetter*[1,3], *Alexandre Tartakovsky*[1,4], *Chung Shih*[1,2]**

[1]*National Energy Technology Laboratory, 3610 Collins Ferry Road, Morgantown, WV 26505, USA*

[2]*NETL Support Contractor, 3610 Collins Ferry Road, Morgantown, WV 26505, USA*

[3]*Battelle, 505 King Avenue, Columbus, OH 43201, USA*

[4]*UIUC Department of Civil and Environmental Engineering, Newmark Civil Engineering Laboratory, MC-250, 205 North Mathews Avenue, Urbana, IL 61801, USA*

Overview of the SMART Initiative

Machine Learning And Advanced Computation Are The Tools We Use

NATIONAL ENERGY TECHNOLOGY LABORATORY

- **The goal of the SMART Initiative is to accelerate decision-making in subsurface applications**
- **Machine learning (ML) is the focus, but in general any advanced computational approaches are being considered**
- **Phase 1 of SMART was research focused, identifying fast, physics-informed methods for several key areas in carbon capture, utilization and storage (CCUS):**
  - Geological characterization from pore-scale to field-scale
  - Fast predictive modeling of subsurface multi-phase physics
  - History matching and site optimization
  - User interaction with the models through a virtual learning platform
- **Phase 2 is focused on *implementing* these techniques for workflows at actual carbon storage sites**

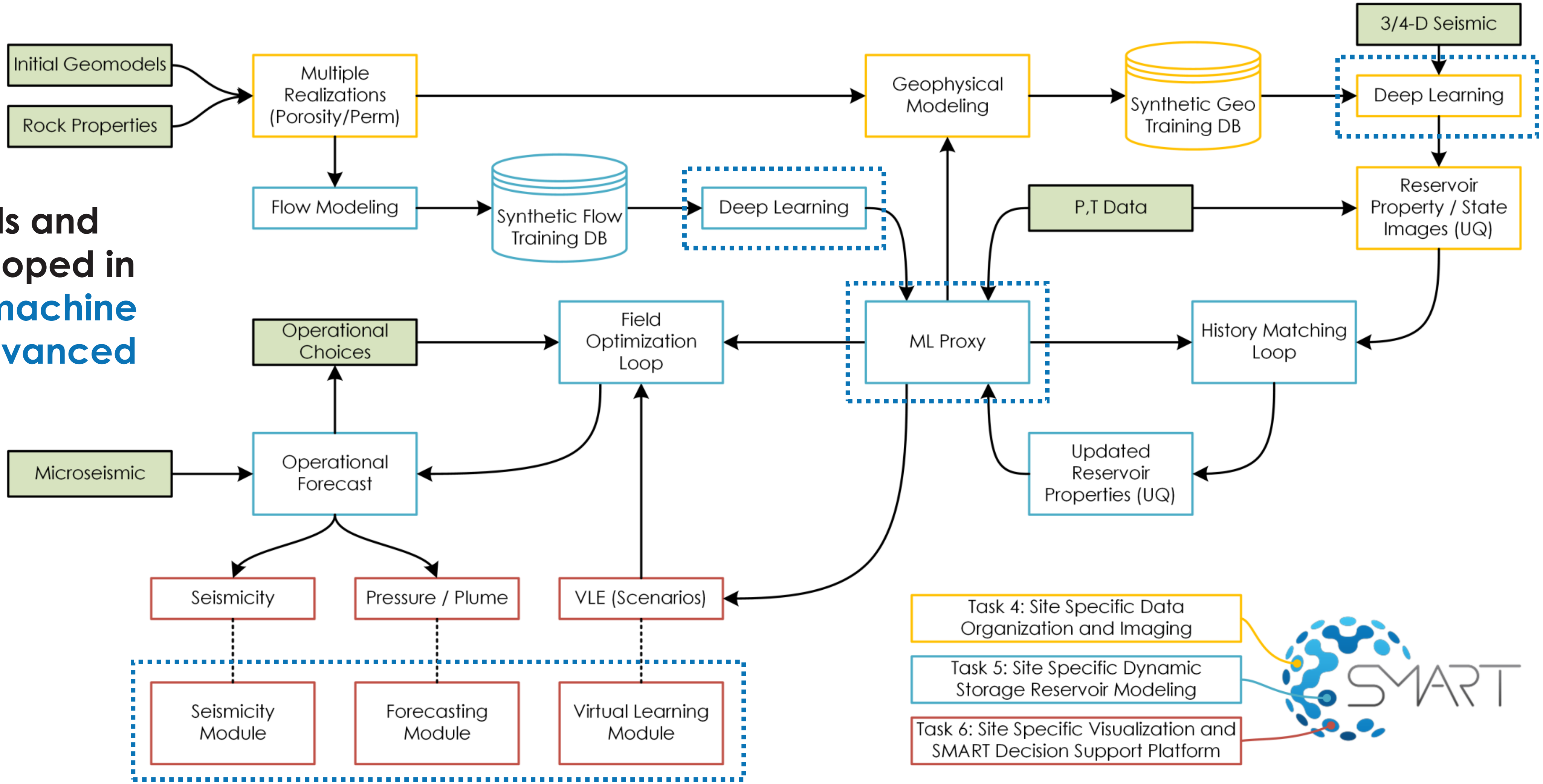
SMART

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

SMART Initiative Phase 2

Areas of Exploration Throughout the Project

Many of the tools and workflows developed in SMART involve machine learning and advanced computation



Task 3: Advanced Machine Learning and Computational Methods

Method and Tool Development to Support SMART Workflows

- **Goals and Objectives:**
  - Identify needs or gaps in site-specific workflows for which SMART Phase 1 solutions are not readily available
  - Develop or update Phase 1 methods to meet these needs
  - Publish tools for implementing these new methods in operational settings
- **Work Completed to Date:**
  - Identified 31 needs for site-specific workflows
  - Grouped these needs into topic areas and selected three topics for investigation
  - Developed preliminary solutions for these topic areas
  - Currently working to formalize and publish many of these tools so other SMART team members can begin using them within their site-specific workflows
  - Also continuing to develop solutions that are at a lower state of maturity

Projects Currently Underway

Topic Area #1: Fast and Flexible Solutions for Fluid Flow Prediction

NATIONAL ENERGY TECHNOLOGY LABORATORY

- **Project Team:** NETL, LLNL, PNNL, SNL, UIUC
- **Objectives**
  - Build flexible models that leverage advanced approaches (e.g., Neural Operators) to handle the dynamic evolution of pressure, saturation, and stress and can serve as the basis to expand to solve other field prediction problems
  - These models also allow for the incorporation of physics and scientific knowledge to increase the user confidence and understanding of the model reasoning processes

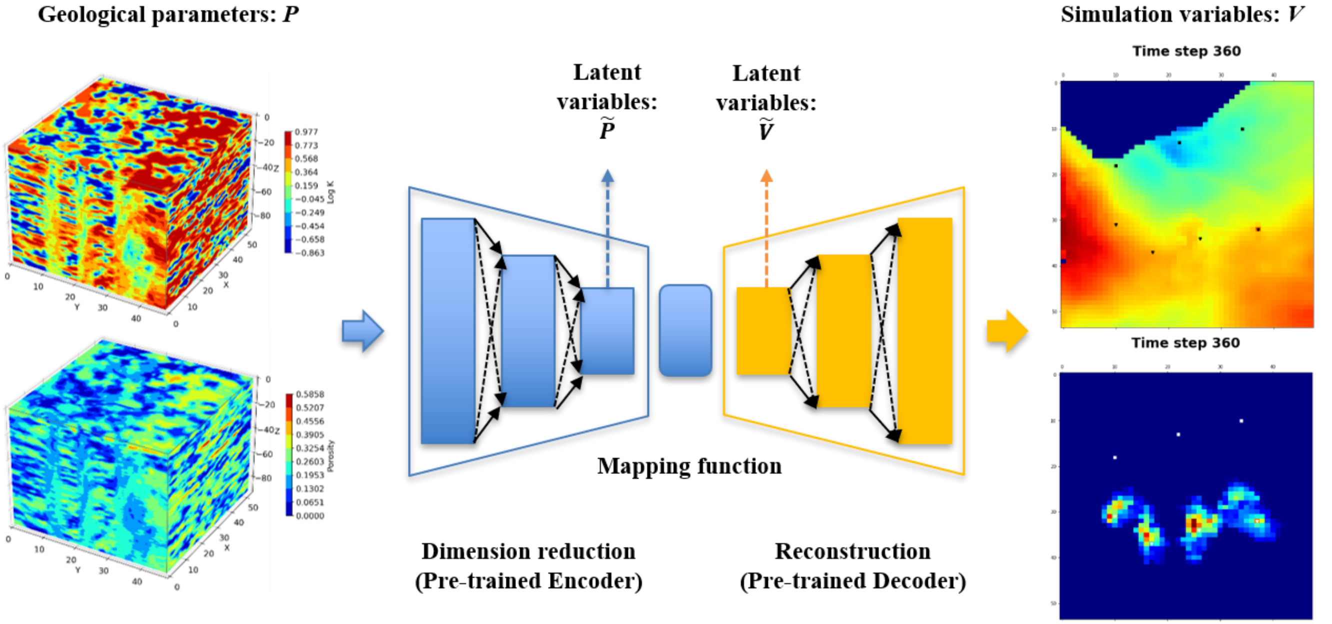
Approach	Type	Organization	POC
DeepONet	ML	PNNL	Amanda Howard
U-Shaped FNO	ML	LLNL	Qingkai Kong
GraphNO	ML	NETL	Chung Shih
PICKLE	ML	UIUC	Alex Tartakovsky
HGGNN	ML	SNL	Meen Teeratorn
Wafer Scale Engine Field Equation Application Programming Interface (WFA)	HPC	NETL	Chung Shih

- **Progress**
  - Initial versions of these approaches were developed and tested on one reservoir (clastic shelf) and are now being adapted for use at the Illinois Basin Decatur Project (IBDP), the main site being focused on in Phase 2

Projects Currently Underway

Topic Area #2: Reduced Dimensional Representations

- **Project Team:** UTBEG, LLNL, SNL, UIUC
- **Objectives**
  - Map complex 3D geological parameters (e.g., porosity, permeability) and corresponding simulation variables (e.g., pressure fields, CO_2 saturation distribution, and CO_2 plume location) to a set of uncorrelated low-dimensional representations (i.e., latent variables)
  - Perform ML within that latent space to model reservoir behavior more efficiently and reliably than ML-based forward models trained in the original parameter space



- **Progress**
  - Trained and evaluated two different mapping approaches from geological parameters to a latent space
  - Developed uncertainty quantification methods
  - Applied the models to a SMART Phase 1, 120-run simulation dataset at the Gulf of Mexico High Island 24L site

Projects Currently Underway

Topic Area #3: Transfer Learning

- **Project Team:** NETL, Battelle, LANL, PNNL, PSU, SNL, TAMU, UTBEG, UU
- **Objectives**
  - As carbon injection and storage becomes more widespread, there will be a need to re-use models and/or make do with as few reservoir simulations as possible
  - The goal on this project is to develop strategies to permit transfer of pre-trained ML models to new scenarios
- **Progress**
  - Identified sets of scenarios in which transfer learning may be needed (right)
  - Built simulation datasets containing pairs of scenarios for one set of conditions vs. another
  - Currently building and testing solutions that allow transfer to occur between the scenarios

Aim	Original Training Scenario	Target Model Task
Transfer Learning Across Operational Conditions	Fixed well locations	New well locations within the same site
	Fixed injection rate	Dynamic injection rate at the same site
Transfer Learning Across Physical Systems	Operational and geological variation within one site	Pressure and saturation at a new site
	Coarse-scale simulations	Fine-scale prediction at the same site
Transfer Learning Across Dynamic Systems	Simpler physics (e.g., single-phase flow)	Complex physics (e.g., multi-phase flow)

Advanced Foundational and Flexible Methods for Fast and High-Fidelity Fluid Flow Predictions

Authors and Contact Information

Chung Yan Shih[1,2], Paul Holcomb[3,4], Amanda Howard[5], Qingkai Kong[6], Hewei Tang[6]

[1]National Energy Technology Laboratory, 3610 Collins Ferry Road, Morgantown, WV 26505, USA

[2]NETL Support Contractor, 3610 Collins Ferry Road, Morgantown, WV 26505, USA

[3]National Energy Technology Laboratory, 626 Cochran Mill Road, Pittsburgh, PA 15236, USA

[4]NETL Support Contractor, 626 Cochran Mill Road, Pittsburgh, PA 15236, USA

[5]Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99354, USA

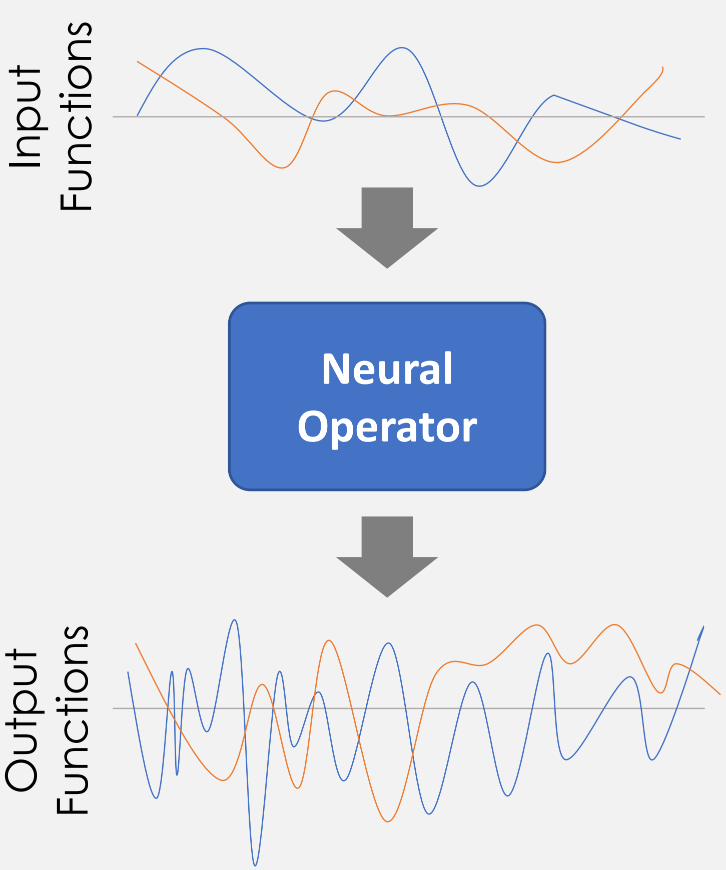
[6]Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94550-9234, USA

The Problem – Need Fast and Flexible Methods

- While many ML models have been developed for subsurface flow, they generally focus on one location or configuration

- There is a need for more flexible models with the same predictive power

Map infinite-dimensional function space



STEPS

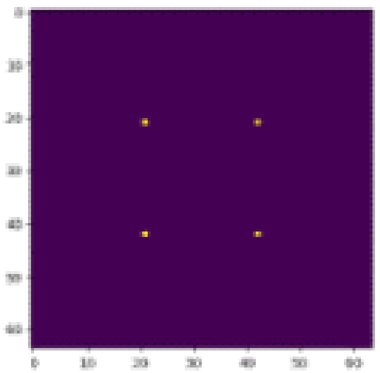
Explore and develop operator learning approach
COMPLETED

Evaluate prediction results
COMPLETED

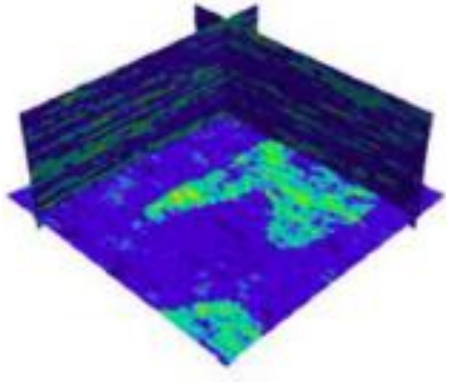
Test the generalizability (Universal approximation, different configurations, etc.)
ONGOING

The Clastic Shelf Dataset

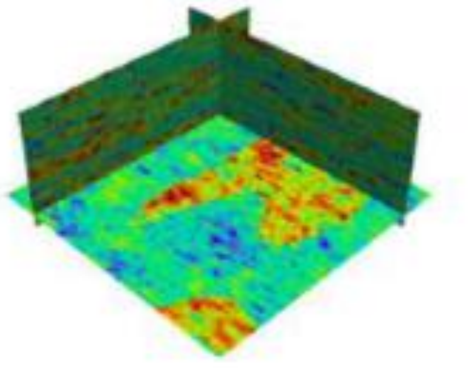
- Clastic Shelf reference reservoir model developed by Tang et al.[1]
  - 2,928 realizations (32,156 x 32,156 x 85 m^3)
  - Data shape: 64 x 64 x 28 (x, y, z)
  - Four injectors equally spaced with an injection rate of 2M metric tons/year over 10 years (1 year Δt)
- Features of the data



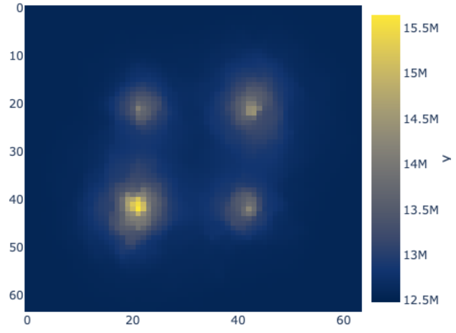
Grid



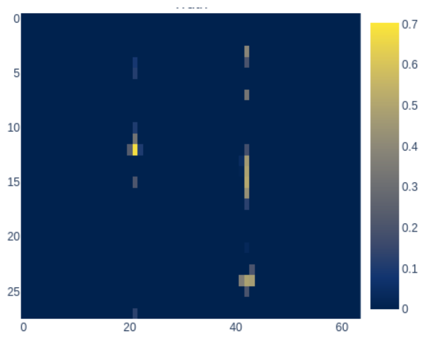
Porosity



Permeability



Pressure



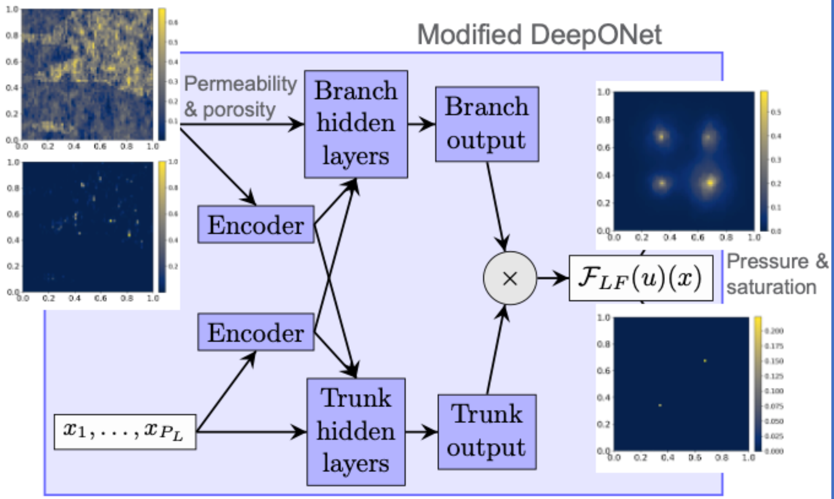
Saturation

[1]Tang et al. “Deep learning-accelerated 3D carbon storage reservoir pressure forecasting based on data assimilation using surface displacement from InSAR” (2022)

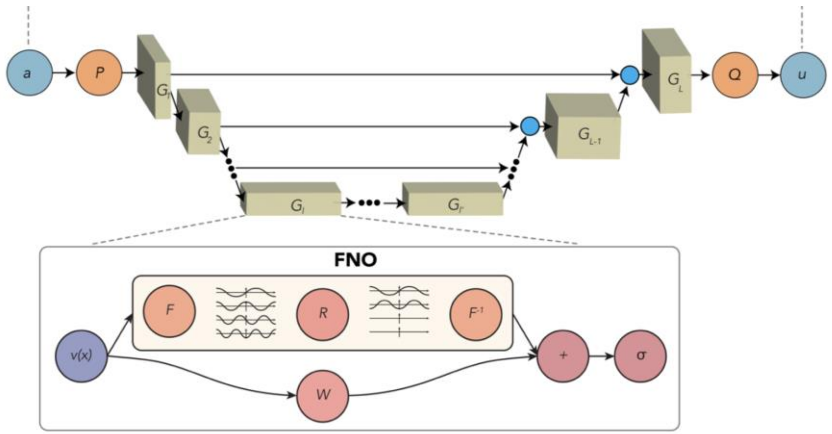
Operator Learning Approaches

Develop Three Advanced Methods for Modeling Subsurface Properties Over Time

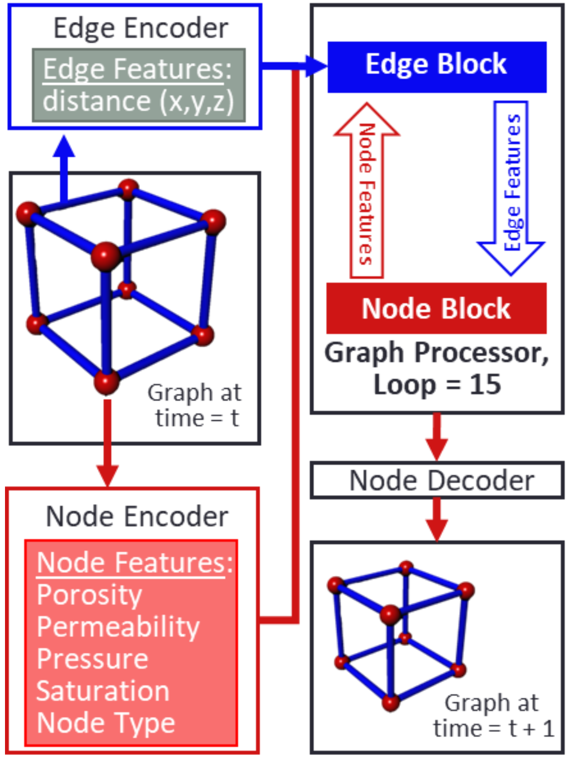
Deep Operator Networks (DeepONets)



U-shaped Fourier Neural Operator (U-NO)



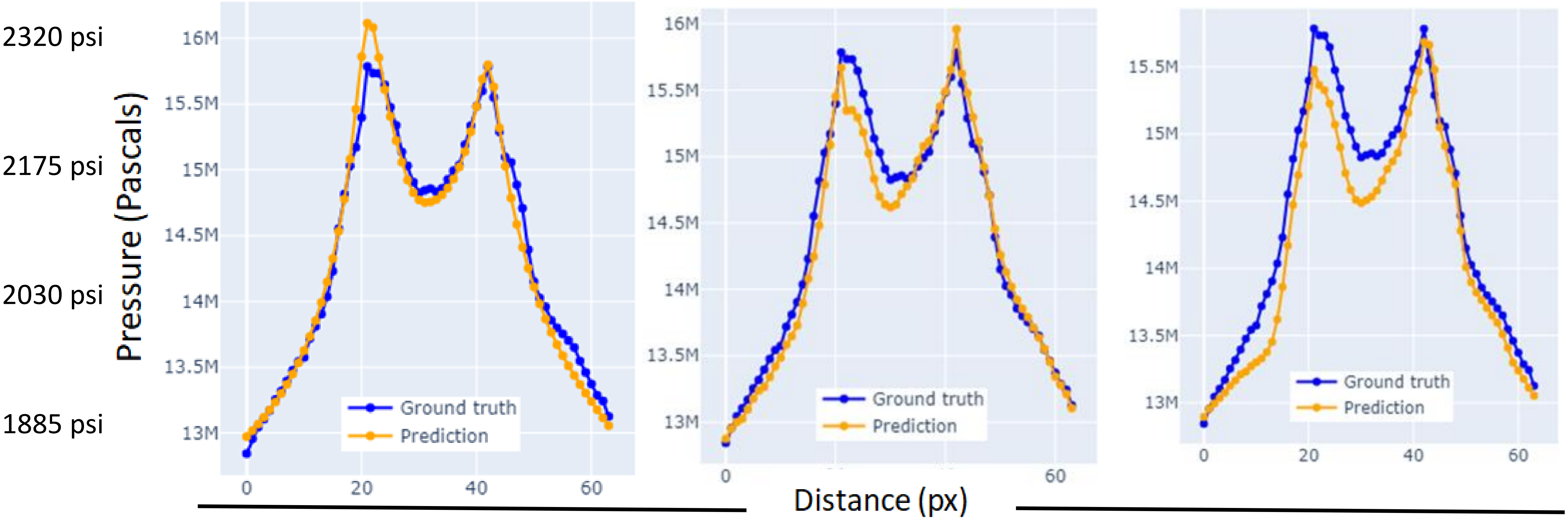
Graph Neural Operator (GNO)



Pressure Profile Through Two Wells

NATIONAL ENERGY TECHNOLOGY LABORATORY

Test Case 1, t = 10 years, y=21, z=2



DeepONets **U-NO** **GNO**

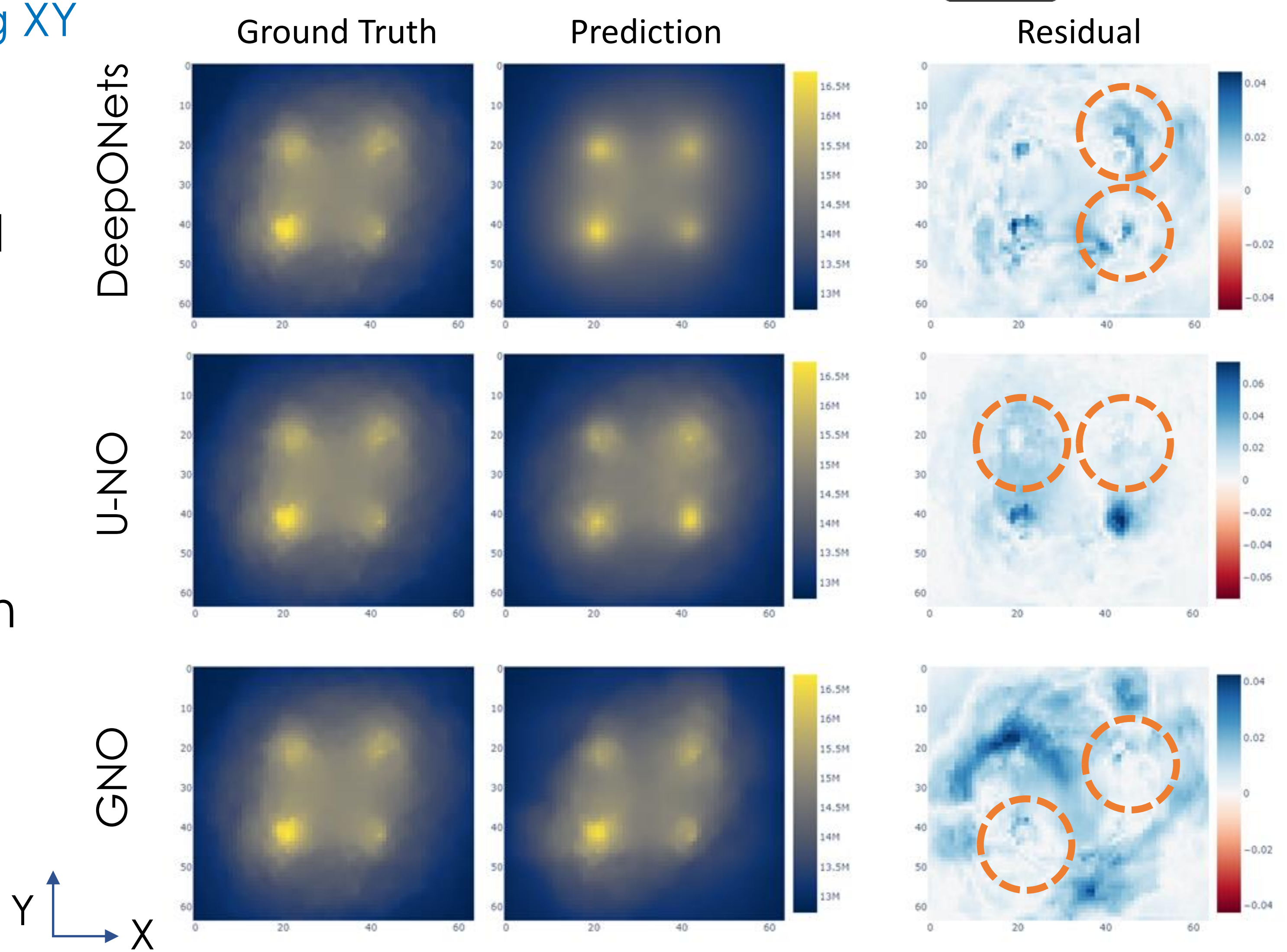
Pressure Prediction Comparison to Ground Truth

NATIONAL ENERGY TECHNOLOGY LABORATORY

Test Case 1, t=10, y=21, z=2 showing XY

- Three approaches performed well on pressure prediction

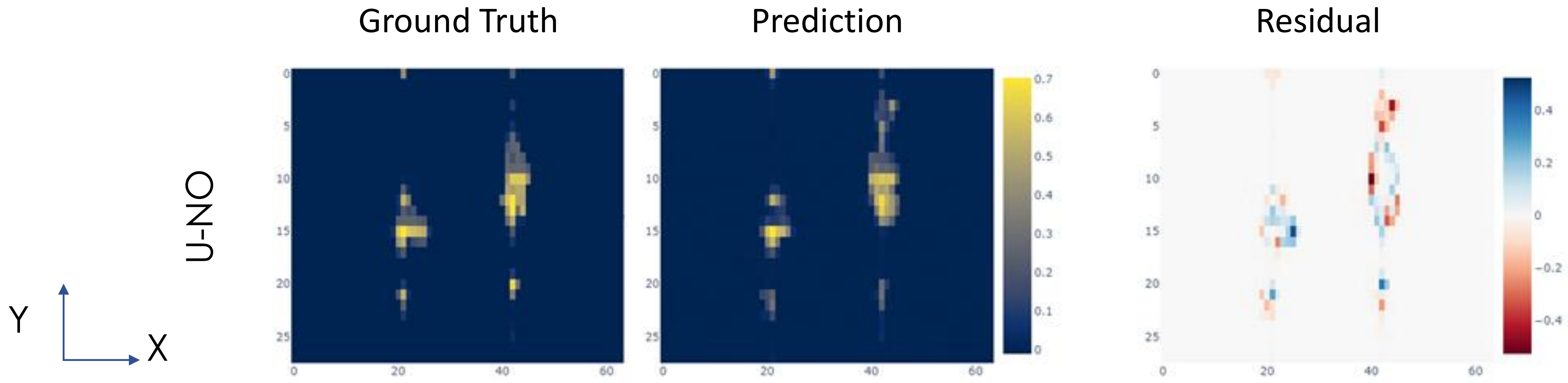
- The different models work better in different regions, indicating that a model ensemble may do better than any of the individual models



Saturation Prediction Comparison to Ground Truth

Test Case 1, t=10, y=21, z=2 showing XZ

- Saturation is harder to predict than pressure due to large amount of zero points and not a lot of variation
- One of the models, U-NO, had reasonable saturation performance (see below), and we are currently investigating issues with other models' predictions



Summary

- Three methods were able to predict pressure with performance similar to other ML methods
- Different model architectures complement each other in prediction results
- The next step is to test operator learning's generalizability by testing different operation conditions

Questions?

NETL RESOURCES

VISIT US AT: **www.NETL.DOE.gov**

@NETL_DOE

@NETL_DOE

@NationalEnergyTechnologyLaboratory