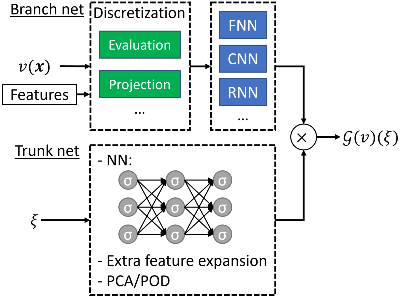
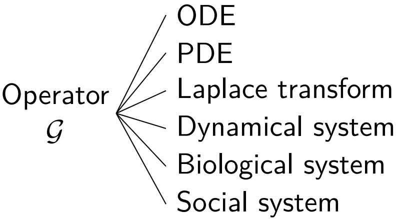
Deep neural operators with **reliable extrapolation** for **multiphysics, multiscale & multifidelity problems**

$$\mathcal{G}(u)(y) \approx \underbrace{\sum_{k=1}^{p} \sum_{i=1}^{n} c_i^k \sigma \left(\sum_{j=1}^{m} \xi_{ij}^k u(x_j) + \theta_i^k \right)}_{\text{branch net}} \underbrace{\sigma(w_k \cdot y + \zeta_k)}_{\text{trunk net}}$$





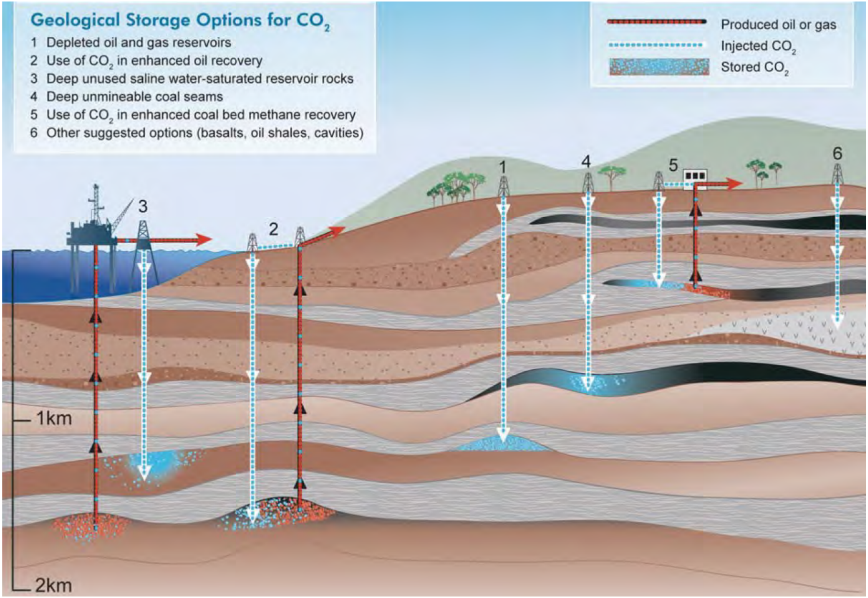
Lu Lu

Department of Chemical and Biomolecular Engineering

Applied Mathematics and Computational Science

University of Pennsylvania

Geological carbon sequestration



Funnell, et al., *GNS Science Report*, 2008

Modeling of geological carbon sequestration

Multiphase flow in porous media ($\alpha = CO_2$ or brine)

$$\frac{\partial M^\alpha}{\partial t} = -\nabla \cdot (\mathbf{F}^\alpha|_{adv} + \mathbf{F}^\alpha|_{dif}) + q^\alpha$$

- Mass: $M^\alpha = \phi \sum_p S_p \rho_p X_p^\alpha$
- Advective mass flux: $\mathbf{F}^\alpha|_{adv} = \sum_p X_p^\alpha \rho_p \mathbf{u}_p$
- Darcy velocity: $\mathbf{u}_p = -k \left(\nabla P_p - \rho_p \mathbf{g}\right) k_{rp}/\mu_p$

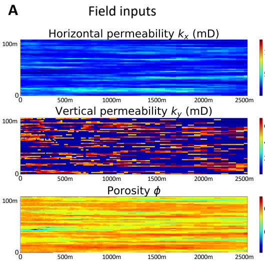
- ϕ: Porosity
- S_p: Saturation
- ρ_p: Density
- X_p^α: Mass fraction
- P_p: Fluid pressure
- μ_p: Viscosity
- $\mathbf{g}$: Gravitational acceleration
- k: Absolute permeability
- k_{rp}: Relative permeability

Challenge: Numerical simulation is computationally expensive.

- Multiphysics & Multiscale
- Large spatial scale (12.5m–200m $\times$ 1,000,000m) & temporal scale (30 years)

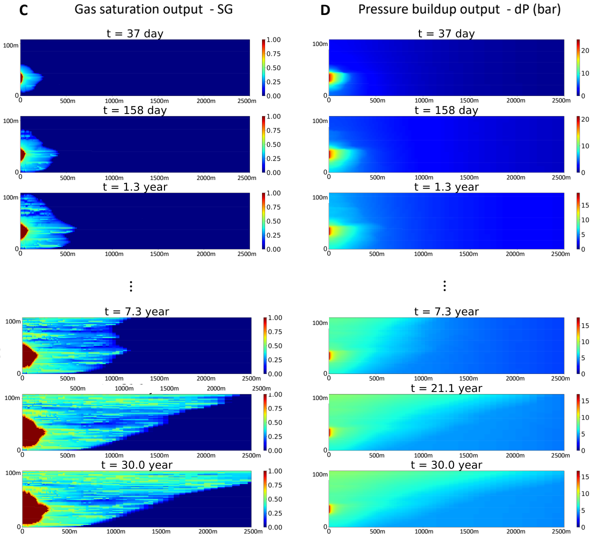
Our approach: Surrogate modeling via machine learning to enable fast prediction

Data example



B Scalar inputs

- Injection rate Q = 1.92 MT/yr
- Iso-thermal reservoir temperature T = 121.2 ℃
- Initial Pressure P_{init} = 258.3 bar
- Irreducible water saturation S_{wi} = 0.28
- Van Genuchten scaling factor λ = 0.45
- Perforation top location $Perf_{top}$ = 32 m
- Perforation bottom location $Perf_{bottom}$ = 38 m



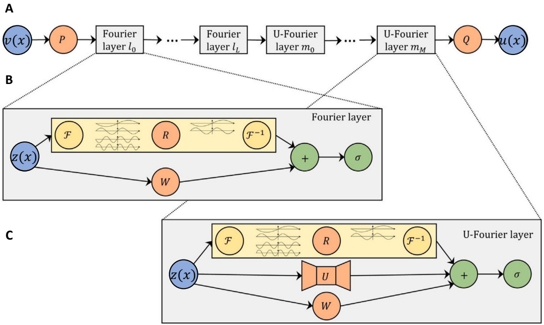
Wen et al., *Adv Water Resour*, 2022

Horizontal axis is truncated for better visualization.

Standard networks

Aim: *Discrete* output in 2D space & 1D time (3D)

- Convolutional neural network (CNN): 3D U-Net
- Fourier neural operator (FNO): 3D FNO
  - Learning in the Fourier space



U-FNO: 3D U-Net + 3D FNO

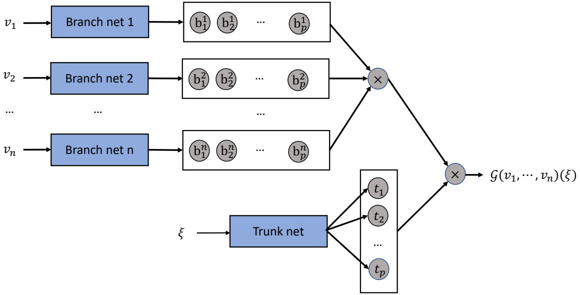
- Good prediction accuracy
- High computational cost

Wen et al., *Adv Water Resour*, 2022

Multiple-input deep operator network (MIONet)

Idea: *Continuous* output in space & time

- Output is a scalar function of $\xi = (x, y, t)$

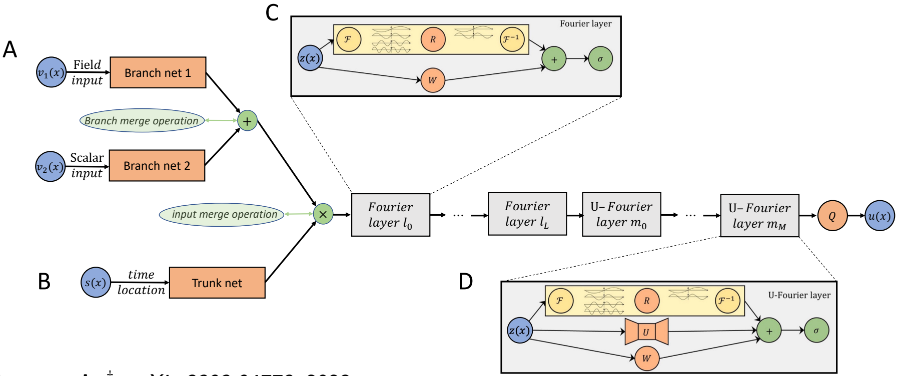


- Low computational cost
- Hard to learn detailed structure in space

Jin, Meng, **Lu**[†], *SIAM J Sci Comput*, 2022

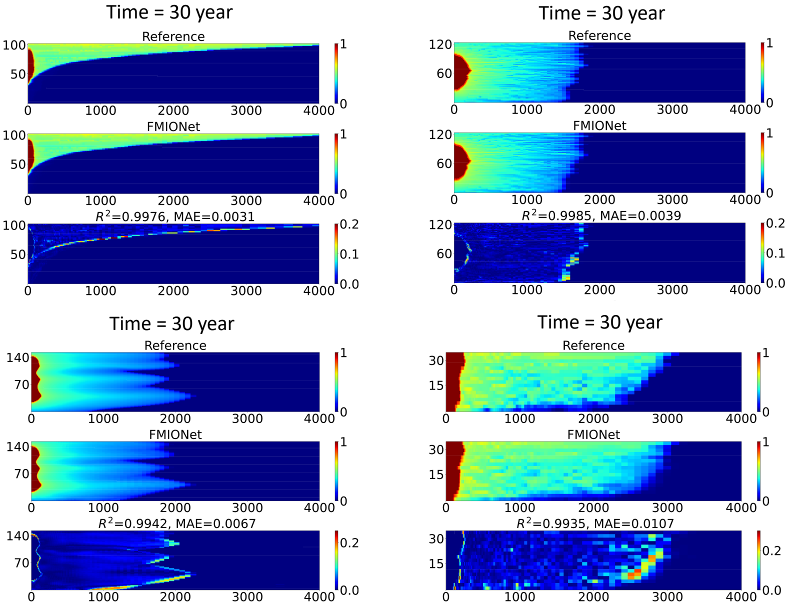
Fourier-MIONet

- Standard: U-FNO (3D U-Net + 3D FNO)
  - Accurate, Expensive
- MIONet
  - Efficient, Hard to learn detailed structure in space
- **Fourier-MIONet**: MIONet + U-FNO
  - Time: Trunk net input
  - Space: *2D* U-FNO as the decoder ("Merge net")



Jiang, ..., **Lu**†, *arXiv:2303.04778*, 2023

Prediction: Gas saturation



Jiang, ..., **Lu**†, *arXiv:2303.04778*, 2023

Fourier-MIONet vs. U-FNO

- **Accuracy**: Almost the same

	R^2	MAE
U-FNO	0.992	0.0031
FMIONet	0.987	0.0033

- **Training**: Much less resources

	# Parameters	CPU memory (GiB)	GPU memory (GiB)	Time (hours)
U-FNO	33,097,829	103	15.9	42.6
FMIONet	3,685,325	15	5.6	12.3

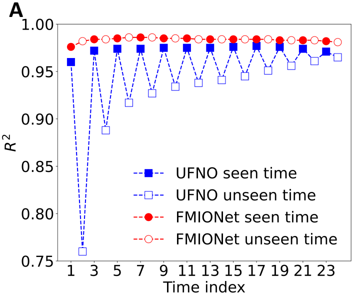
- **Prediction**: Much less resources

	CPU memory (GiB)	GPU memory (GiB)	Time (s)
U-FNO	15.3	7.1	0.075
FMIONet	5.1	3.5	0.041

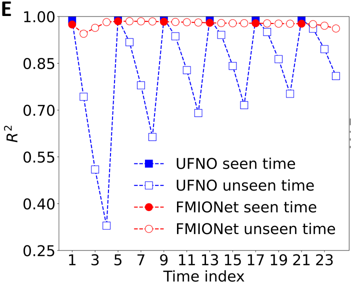
Jiang, ..., **Lu**†, *arXiv:2303.04778*, 2023

Prediction for *unseen* time

50% training data



25% training data

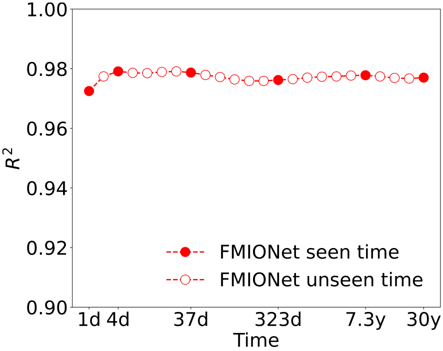


Good generalization even for *unseen* time!

- Fourier-MIONet obeys physics: Continuity over time.

Jiang, ..., **Lu**†, *arXiv:2303.04778*, 2023

Nonuniform sampling of training data



$R^2 > 0.97$ with only 6 different time data for training

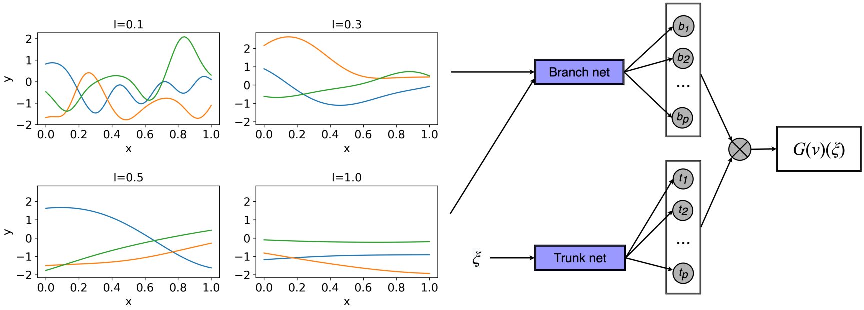
Machine learning models, including DeepONets, are limited to interpolation.

Extrapolation?

Operator learning extrapolation

Learn an operator $\mathcal{G} : v(x) \mapsto u(\xi)$

- Gaussian random field (GRF): $v(x) \sim \mathcal{GP}(0, k_l(x_1, x_2))$
- Radial-basis function (RBF) kernel: $k_l(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2l^2}\right)$
- l: Correlation length



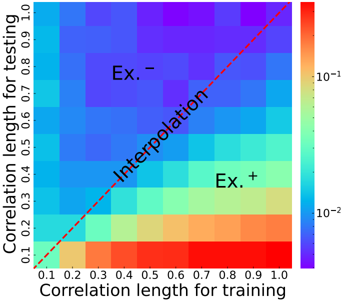
Zhu, ..., **Lu**[†], *Comput Methods Appl Mech Eng*, 2023

Extrapolation examples

An ODE ($x \in [0,1]$)

$$\frac{du}{dx} = v(x), \quad u(0) = 0$$

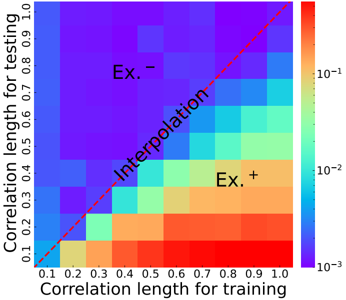
$$\mathcal{G} : v(x) \mapsto u(x)$$



Diffusion-reaction equation ($(x,t) \in [0,1]^2$)

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ku^2 + v(x)$$

with zero IC/BC, $D = 0.01$, $k = 0.01$



$$\text{Prediction} = \begin{cases} \text{In.} & \text{when } l_{\text{train}} = l_{\text{test}} \\ \text{Ex.}^- & \text{when } l_{\text{train}} < l_{\text{test}} \\ \text{Ex.}^+ & \text{when } l_{\text{train}} > l_{\text{test}} \end{cases}$$

Zhu, ..., **Lu**†, *Comput Methods Appl Mech Eng*, 2023

Quantify extrapolation complexity

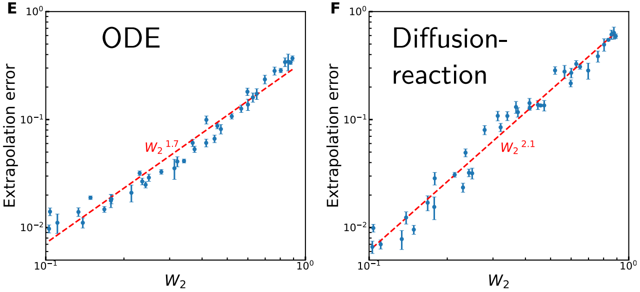
Two GRFs: $f_1 \sim \mathcal{GP}(m_1, k_1)$, $f_2 \sim \mathcal{GP}(m_2, k_2)$

2-Wasserstein (W_2) metric: Distance between two spaces

$$W_2(f_1, f_2) := \left\{ \|m_1 - m_2\|_2^2 + \mathrm{Tr}\left[K_1 + K_2 - 2\left(K_1^{\frac{1}{2}} K_2 K_1^{\frac{1}{2}} \right)^{\frac{1}{2}} \right] \right\}^{\frac{1}{2}}$$

where $K_i : L^2(X) \to L^2(X)$ is the covariance operator of k_i

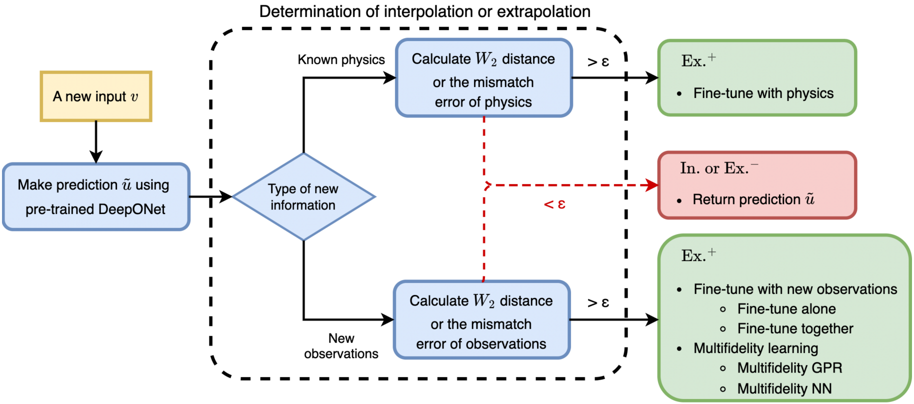
$$[K_i \phi](x) = \int_X k_i(x, s)\phi(s)ds, \quad \forall \phi \in L^2(X)$$



$\log$ Extrapolation error $\propto \log W_2$

Zhu, ..., **Lu**†, *Comput Methods Appl Mech Eng*, 2023

Reliable extrapolation



- In. or Ex.$^-$: Return prediction $\tilde{u}$
- Ex.$^+$: Additional information to correct $\tilde{u}$ (fine-tune or multifidelity learning)
  - **Physics**: Governing PDEs
  - New **data** at sparse locations (high-fidelity)

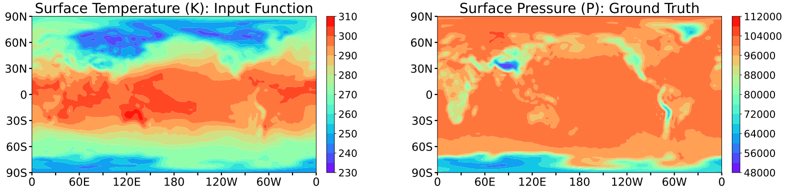
Zhu, ..., **Lu**[†], *Comput Methods Appl Mech Eng*, 2023

Global climate change

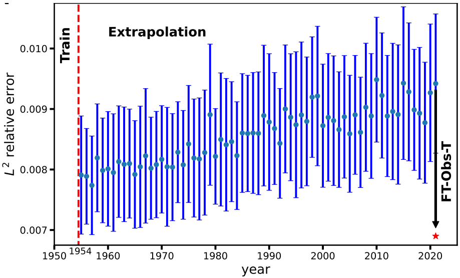
Daily surface air temperature $T(\mathbf{x})$ & pressure $p(\mathbf{x})$ from 1950 to 2021

NCEP-NCAR Reanalysis Database

$$\mathcal{G} : T(\mathbf{x}) \mapsto p(\mathbf{x})$$



January 1, 2021

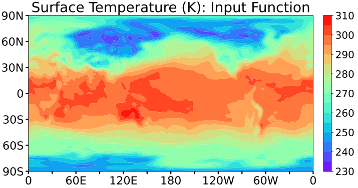


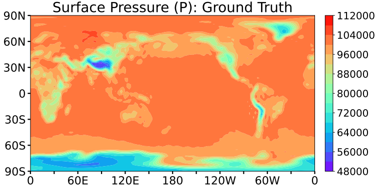
- Train: 1950–1954
- For later years, larger extrapolation error

Zhu, ..., **Lu**†, *Unpublished*

Global climate change

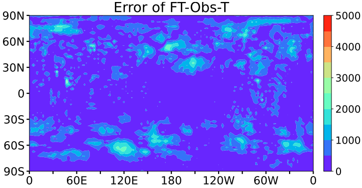
$$\mathcal{G} : T(\mathbf{x}) \mapsto p(\mathbf{x})$$

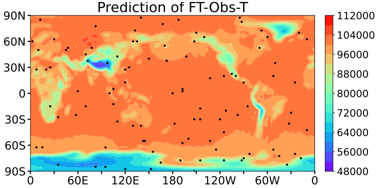




January 1, 2021

Data: 100 weather stations





L^2 relative error: 0.69%

Zhu, ..., **Lu**†, *Unpublished*

DeepONets

- **A family of DeepONets**
  - DeepONet (**Lu** et al., *Nature Mach Intell*, 2021)
  - MIONet: Multiple-input operator (Jin, Meng, **Lu**†, *SIAM J Sci Comput*, 2022)
  - POD-DeepONet (**Lu** et al., *Comput Methods Appl Mech Eng*, 2022)
  - Fourier-DeepONet/MIONet (Jiang, ..., **Lu**†, *arXiv:2303.04778*, 2023; Zhu, ..., **Lu**†, arXiv:2305.17289)
  - DeepM&Mnet (Cai, Wang, **Lu**, et al., *J Comput Phys*, 2021; Mao, **Lu**, et al., *J Comput Phys*, 2021)
  - Multifidelity DeepONet (**Lu**† et al., *Phys Rev Res*, 2022)
  - Reliable extrapolation (Zhu, ..., **Lu**†, *Comput Methods Appl Mech Eng*, 2023)
- **Theory**
  - Universal approximation theorem (Jin, Meng, **Lu**†, *SIAM J Sci Comput*, 2022)
  - Error analysis (Deng, Shin, **Lu**, et al., *Neural Netw*, 2022)

Accuracy **Efficiency** **Capability**

- **Multiphysics & Multiscale applications**
  - High-speed boundary layer (Di Leoni, **Lu**, et al., *J Comput Phys*, 2023)
  - Electroconvection (Cai, Wang, **Lu**, et al., *J Comput Phys*, 2021)
  - Hypersonics (Mao, **Lu**, et al., *J Comput Phys*, 2021)
  - Geological carbon sequestration (Jiang, ..., **Lu**†, *arXiv:2303.04778*, 2023)
  - Full waveform inversion (Zhu, ..., **Lu**†, arXiv:2305.17289)

Open-source software: **DeepXDE**

Scientific machine learning **Physics-informed learning**

>400,000 Downloads
>100 Research papers

>1,600 GitHub Stars
>50 Contributors around the world

- Universities (>70)
- National labs & Research institutes (>15)
- Industry