Adaptive Physics Transformer with Fused Global–Local Attention for Subsurface Energy Systems

Xin Ju[1,4] Nok Hei (Hadrian) Fung[2] Yuyan Zhang[2] Carl Jacquemyn[2] Matthew Jackson[2] Randolph Settgast[3,4] Sally M. Benson[1,4] Gege Wen[2,4]

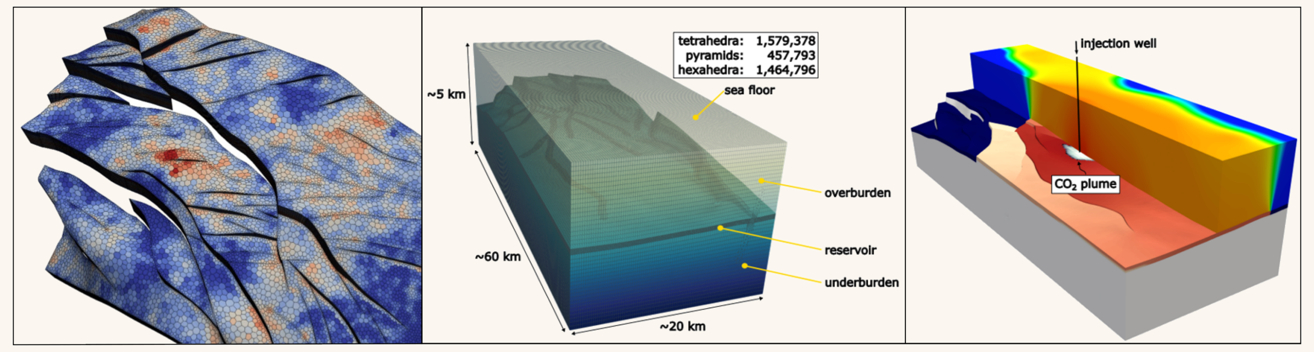
[1] Stanford University [2] Imperial College London [3] Lawrence Livermore National Laboratory [4] EarthFlow AI

EarthFlowAI ICML International Conference On Machine Learning

The Challenge

Full-physics simulation of subsurface systems — CO_2 storage, geothermal, hydrocarbons — is notoriously computationally expensive:

- **Heterogeneity:** permeability spans orders of magnitude within metres
- **Multi-scale coupling:** plumes in metres, pressure in kilometres, horizons in decades
- **Cost:** hours per run; UQ and optimization need thousands of runs



faulted heterogeneous geology basin-scale meshes, millions of cells full-physics CO_2 simulation
Left: A. Mazuyer, Sismage-CIG, B. Wendebourg, and F. Lepage. Right: Settgast et al. (2024), GEOS, J. Open Source Software.

One Model, Any Mesh

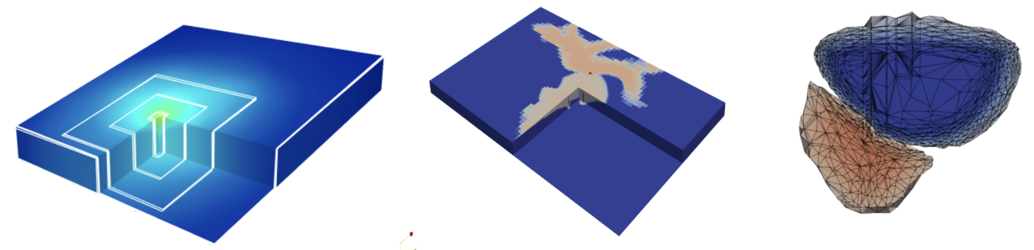
Geological system — any mesh
heterogeneous permeability & porosity wells and schedules
regular, irregular, nested & adaptive meshes

↓

Adaptive Physics Transformer
$\hat{z}(x,t) = \mathcal{D} \circ \mathcal{A}_t \circ \mathcal{E}(a)$
fused local-graph + global-attention neural operator

↓

Physical fields at any query point



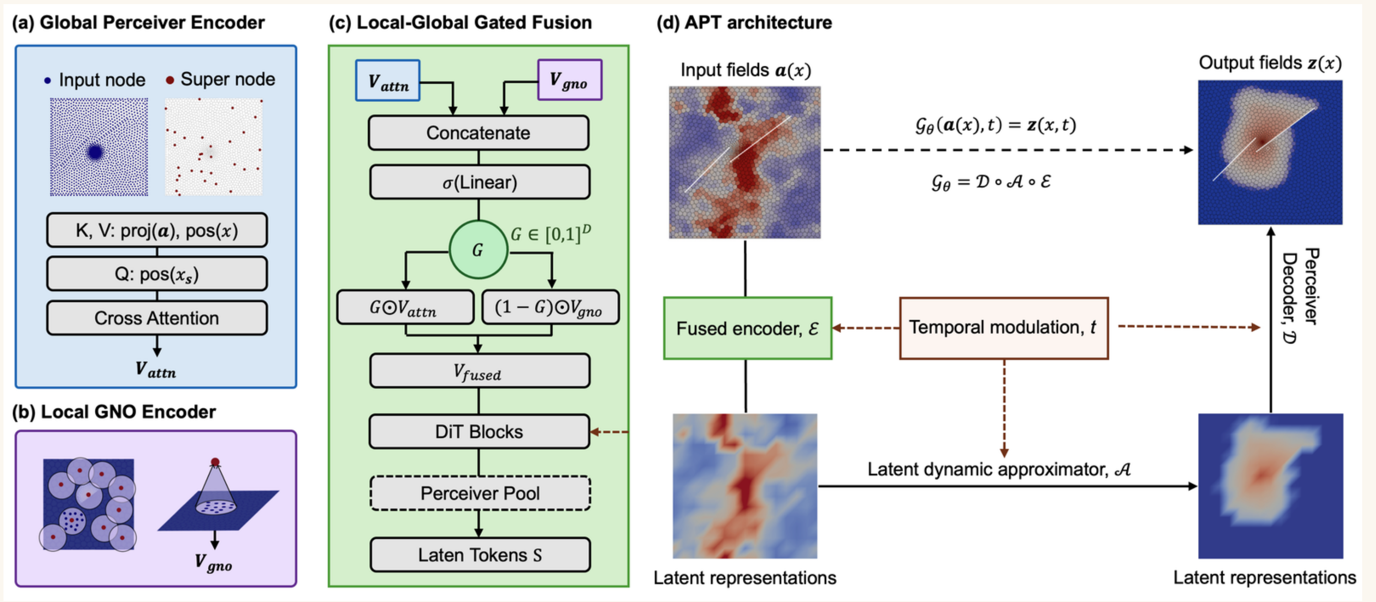
pressure ΔP_g saturation S_g temperature T

Simulate subsurface physics **accurately**, at a **fraction of the cost**, on the meshes engineers actually use.

Contributions

- First neural operator to learn **directly from adaptive-mesh (DMO) simulations**
- **Fused encoder:** local graphs resolve heterogeneity, global attention captures pressure coupling
- **SOTA accuracy** on five subsurface benchmarks, with mesh-convergent super-resolution
- **One scalable backbone** across meshes, physics and datasets — a path to subsurface foundation models

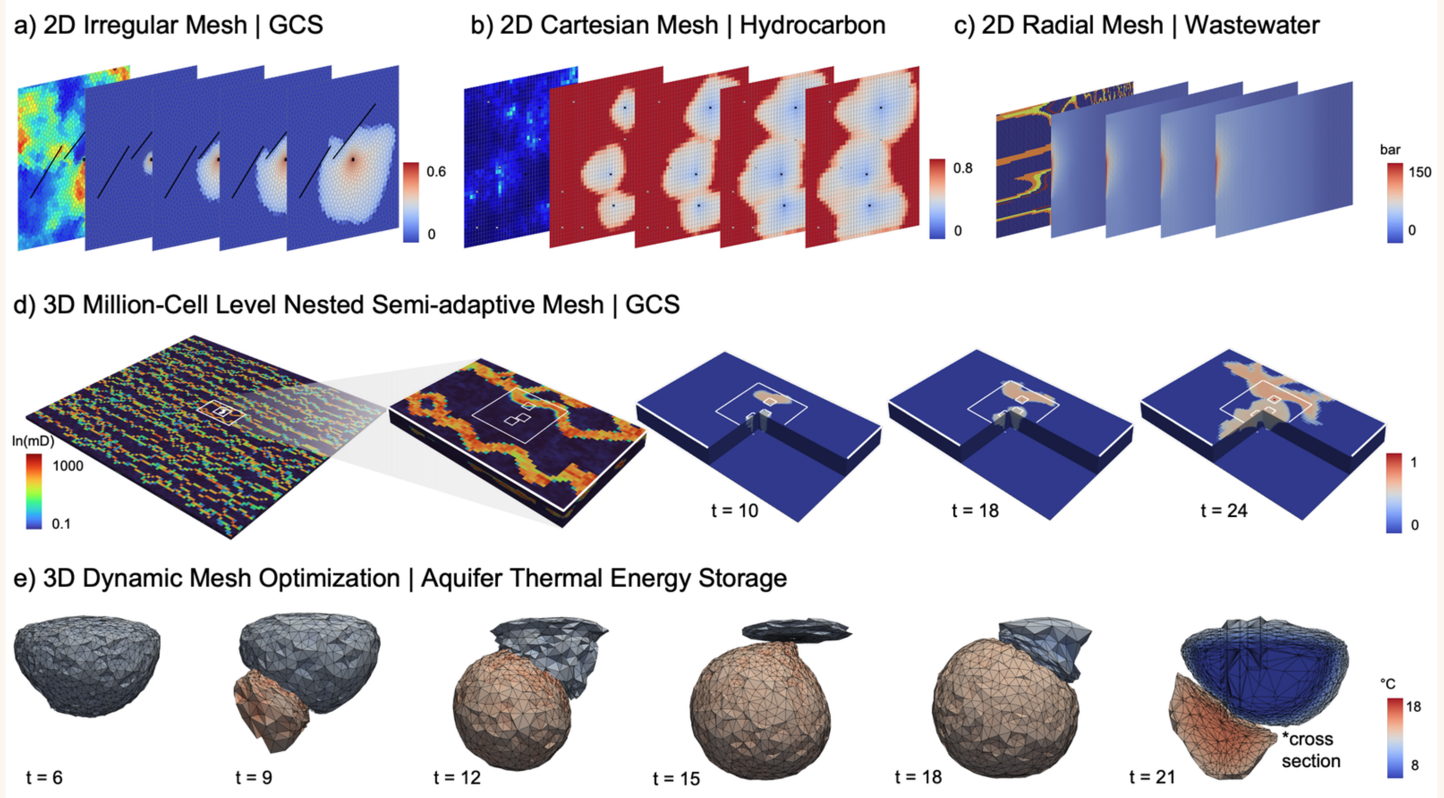
Fused Global–Local Architecture



Fused encoder $\mathcal{E}$ (global + local branches), latent approximator $\mathcal{A}_t$ on fixed-size tokens, perceiver decoder $\mathcal{D}$ at arbitrary coordinates.

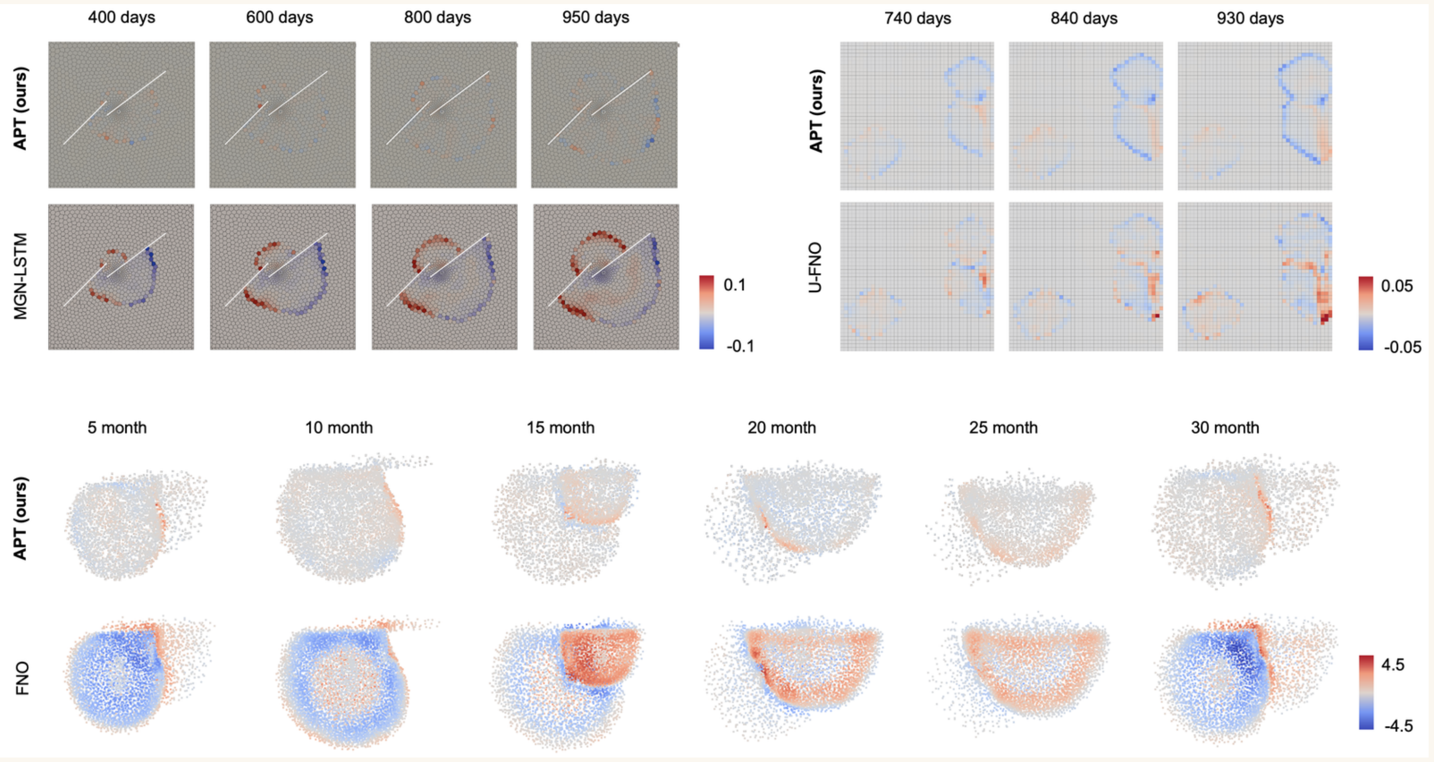
- **Global perceiver encoder** — cross-attention onto fixed supernode queries: long-range physics at $\sim\mathcal{O}(N)$ cost
- **Local GNO encoder** — radius-graph pooling preserves high-frequency property variation at faults and fronts
- **Gated fusion** — learned gate $G \in [0,1]^d$ blends $\mathbf{v}_{\text{attn}}$ and $\mathbf{v}_{\text{gno}}$ per channel
- **Unified modulation** — DiT-style conditioning injects time and operating conditions
- **Mesh-free decoding** — query any (x,t): super-resolution and time interpolation for free

Five Subsurface Benchmarks



Faulted 2D GCS (irregular) hydrocarbon extraction (Cartesian) wastewater injection (radial) basin-scale GCS (nested LGR) ATES (dynamic adaptive mesh)

Accuracy Across Mesh Types



Error maps: APT stays sharp at faults, fronts and plumes.

Faulted 2D CO_2 storage — irregular mesh

Model	$\delta\Delta P_g$ (%)	δS_g (%)	Params
U-FNO	0.34	2.40	1.48M
MGN-LSTM	0.20	1.20	1.38M
Transolver	0.87	3.82	1.59M
UPT	0.43	4.33	1.68M
APT (fused)	**0.11**	**0.32**	1.38M

12× lower saturation error than Transolver.

ATES — adaptive mesh (DMO), first of its kind

APT trains directly on raw, time-varying DMO point clouds.

Model	Rel. L_2	R^2 (%)	Params
FNO (interp.)	0.046	90.6	21M
UPT	0.0401	96.0	19.6M
Transolver	0.0237	98.6	17.5M
MINO	OOM (>80 GB)		–
APT (fused)	**0.011**	**99.0**	17M

4× better than interpolated grids; **7 h → 5 s** per forecast.

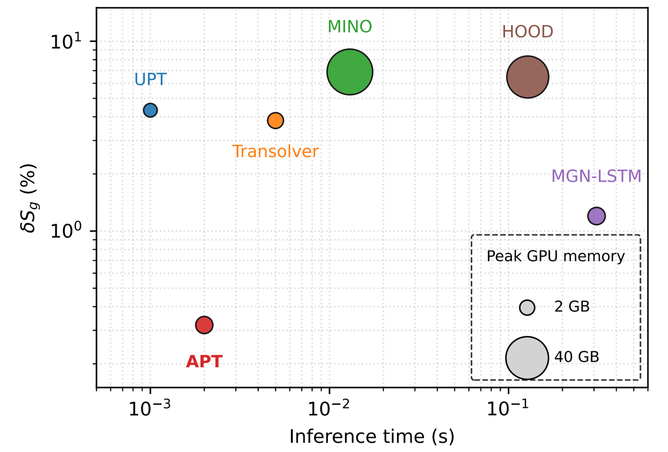
Basin-scale CO_2 storage — nested LGR

0.3–1.0M cells, $60{,}000\times$ cell-volume contrast.

Model	$\delta\Delta P_g$ (%)	δS_g (%)	Params	Models
FNO-DeepONet	0.56	**1.46**	96.3M	5
Nested-FNO	**0.47**	1.79	682.4M	5
APT (fused)	0.56	2.5	**17.9M**	**1**

One unified model — **97%** fewer parameters than Nested-FNO — and **6,800×** faster than the ECLIPSE simulator: full-field inference in **1.4–8 s** instead of 2.75–15.93 h.

Efficiency



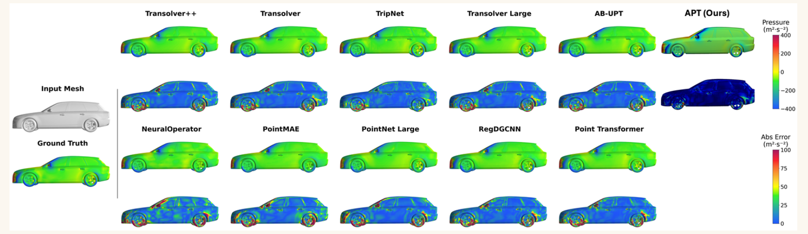
Model	Time (s)	Speedup
GEOS[a]	49.02	1×
MGN-LSTM[b]	0.31	158×
HOOD[c]	0.128	383×
Standard MGN[b]	0.07	700×
MINO[b]	0.013	3,668×
Transolver[b]	0.005	9,593×
APT[b]	**0.002**	**19,873×**
UPT[b]	0.001	44,939×

[a] Intel Xeon E5-2695 v4, single-core serial run. [b] NVIDIA H100 80 GB. [c] NVIDIA A100 80 GB.

19-step rollout (950 days) on faulted 2D CO_2, batch 340: APT leads the accuracy–efficiency Pareto frontier — lowest error at 2 ms per sample, **19,873×** faster than single-core GEOS.

Super-Resolution: General Physics

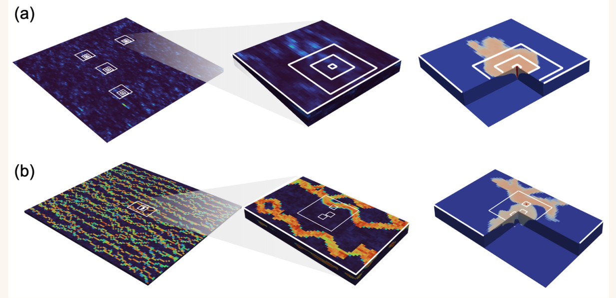
Beyond the subsurface: aerodynamic surface pressure on CarBench. Train on 10k nodes, test on the full ~500k-node car (R^2, %):



Model	10k	Full ~500k
Transolver	96.5	86.6 (↓9.9)
AB-UPT	97.1	86.9 (↓10.2)
APT	96.0	**96.5** (↑**0.5**)

The **only** model that improves at full resolution — a true mesh-convergent continuous operator.

Generalization & Cross-Dataset



(a) Gaussian, 4-level LGR
(b) channelized, 3-level LGR

- **Cross-dataset:** joint training on incompatible LGR meshes lifts channelized R^2 32.0% → **86.7%** (**2.7×**)
- **OOD geology:** best error (**3.28%**) on geostatistical classes never seen in training

Takeaways

- Geometry-, mesh- and physics-agnostic neural operator
- First to learn directly from adaptive-mesh simulation
- SOTA accuracy at a fraction of the parameters & cost
- A scalable backbone for subsurface foundation models

Paper
arXiv.org/abs/2602.11208

LinkedIn
Isaac Ju

Contact
ju1@stanford.edu
isaacju@earthflow.ai

Stanford University IMPERIAL Lawrence Livermore National Laboratory