Learning Large-scale Subsurface Simulations with a Hybrid Graph Network Simulator

Tailin Wu, Qinchen Wang, Yinan Zhang, Rex Ying,
Kaidi Cao, Rok Sosič, Ridwan Jalali, Hassan Hamam,
Marko Maucec, Jure Leskovec

Motivation

Subsurface simulations use computational models to predict the flow of fluids (e.g., oil, water, gas) through porous media. These simulations are pivotal in industrial applications such as petroleum production. Recently, data-driven surrogate models provide a promising complementary approach. However, these models are insufficient in two main ways:

1. Cannot simultaneously model lower-level interactions between neighboring cells and global dynamics such as pressure.
2. Do not scale to large-scale simulations as they have only been applied to 2D grids with up to 10k vertices

Our contributions:

We introduce Hybrid Graph Network Simulator (HGNS), a data-driven surrogate model for learning reservoir simulations of 3D subsurface fluid flows. HGNS is able to:

- Scale to grids with millions of cells per time step, two orders of magnitude higher than previous surrogate models.
- Accurately predict the fluid flow for years into the future.
- Reduce the inference time up to 18 times compared to standard subsurface simulators.
- Outperform other learning-based models by reducing long-term prediction errors by up to 21%.

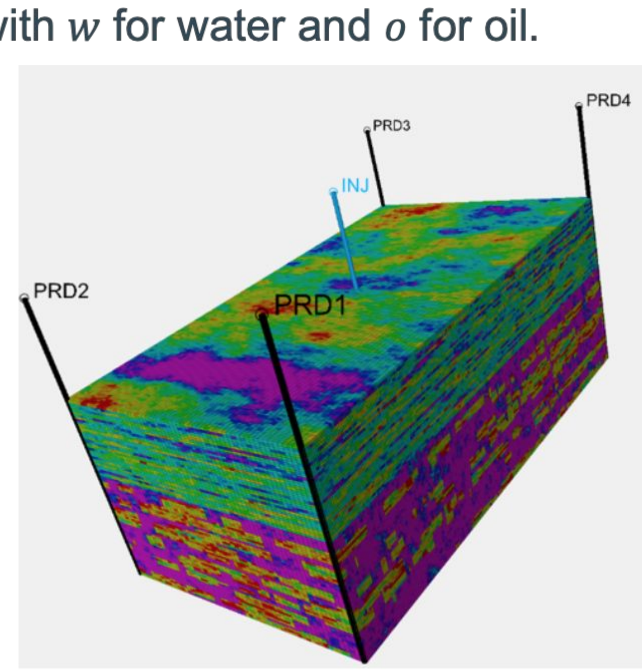
HGNS models how the pressure and saturation of fluids evolve over time, given initial states, static properties of the rock, and external control variables such as injection of water.

Background

We consider the problem of subsurface simulation of oil-water flow, which models how the pressure and saturation of fluids evolve over time, given initial states, static properties of the rock, and external control variables such as injection of water. Here we present a simplified Partial Differential Equation (PDE) for the system:

$$\frac{\partial(\phi \rho_j S_j)}{\partial t} = \nabla \cdot \left(\frac{\rho_j}{\mu_j} k_{rj}(S_j)\mathbf{k}\nabla P\right) + q_j$$

$j = w, o$: different components/phases, with w for water and o for oil.
$S_j \in [0, 1]$: the saturation for phase j
P: pressure
ρ_j: phase density
μ_j: phase viscosity
ϕ: rock porosity
$k_{rj}(S_j)$: relative permeability
k: absolute permeability tensor
q_j: source/sink term



Model architecture and inspiration

Our Hybrid Graph Network Simulator (HGNS) architecture (Fig. 1) consists of a Subsurface Graph Neural Network (SGNN) to model the dynamics of fluids (water, oil) on a finer scale, and a 3D-U-Net (Çiçek et al. (2016)) to model the more global dynamics of pressure. Concretely, our HGNS $f_\theta = (g_\theta, h_\theta)$ can be written as:

$$\begin{cases} \hat{S}^{t+1} = g_\theta(X^t, Q, U^t) + S^t \\ \hat{P}^{t+1} = h_\theta(X^t, Q, U^t) + P^t \end{cases}$$

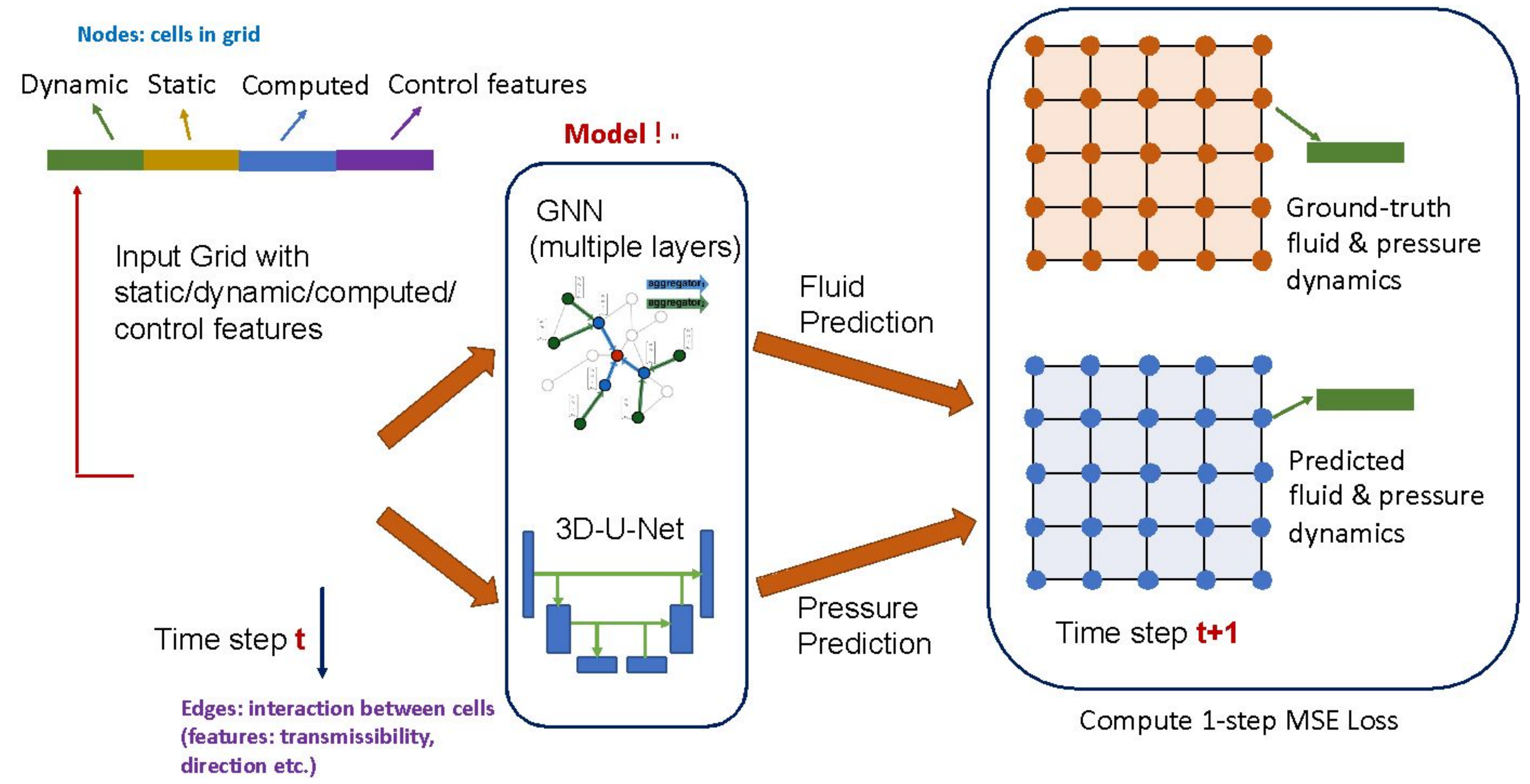
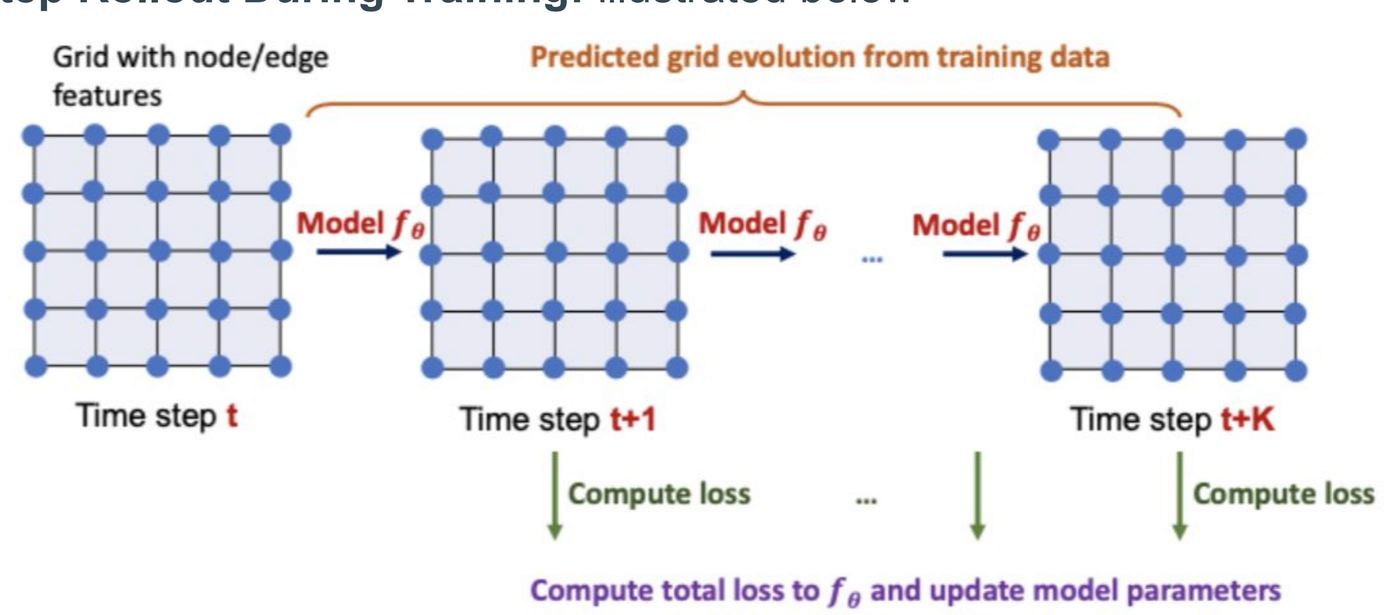


Fig. 1 Overview of our HGNS architecture

Hybrid architecture: Our SGNN g_θ uses the encoder-processor-decoder architecture. The encoder embeds the input into a latent graph.. The processor is a stack of M graph neural network layers, each one performing one round of message passing that mimics the flow of fluid between neighboring cells. The decoder is a simple MLP that maps the output of the processor back to the predicted dynamic variables at the next time step. We use 3D-U-Net (Çiçek et al. (2016)) h_θ to model the more global dynamics of pressure.

Sector-based training: Training of the model presents a special challenge, since realistic data is too large to fit in a single GPU. To address this problem, we developed sector-based training, which partitions data into many sectors, which can be then processed individually.

Multi-step Rollout During Training: illustrated below



Experiments with SPE-10 dataset

Table 1 shows MAE of our HGNS and other baseline learning-based models on 4 unseen trajectories with novel geology and well placements. Table 2 shows ablations:

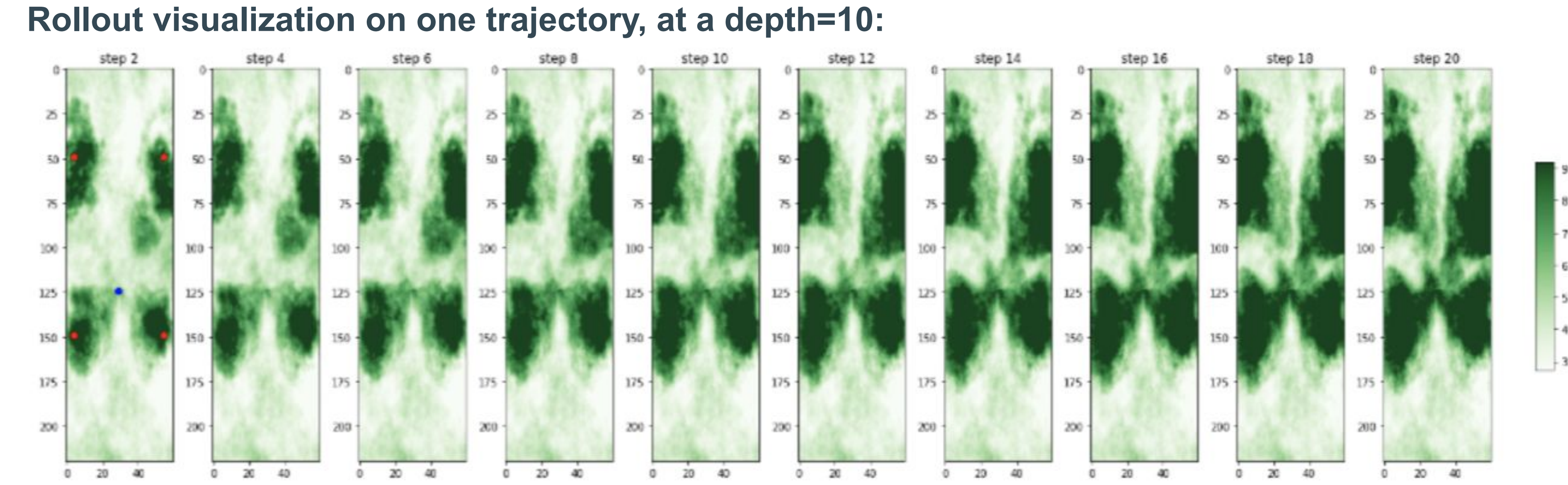
Model	10-step prediction MAE			20-step prediction MAE		
	Pressure (psi)	Water (barrel)	Oil (barrel)	Pressure (psi)	Water (barrel)	Oil (barrel)
Predict no change	210.8	0.941	0.941	296.1	1.541	1.541
CNN	77.9	0.628	0.608	104.2	1.157	1.090
3D-U-Net	94.6	0.361	0.361	142.6	0.725	0.724
HGNS (ours)	**73.6**	**0.286**	**0.286**	**97.2**	**0.664**	**0.664**

Table 1 Comparison with baselines

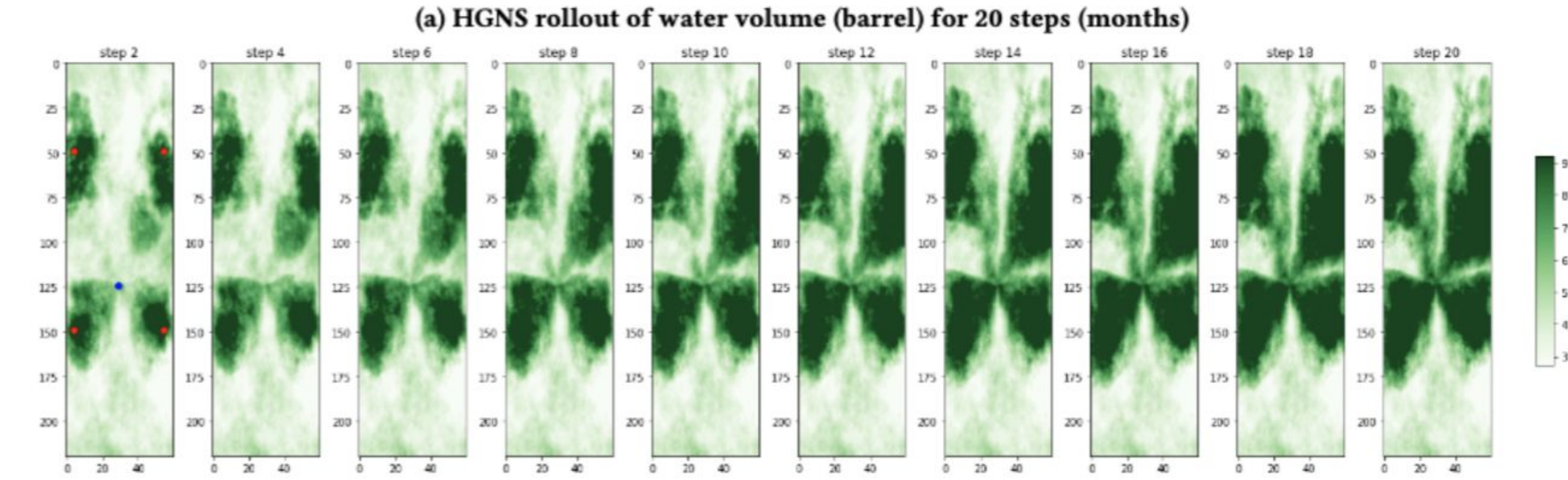
Model	10-step prediction MAE			20-step prediction MAE		
	Pressure (psi)	Water (barrel)	Oil (barrel)	Pressure (psi)	Water (barrel)	Oil (barrel)
HGNS (ours)	73.6	**0.286**	**0.286**	**97.2**	**0.664**	**0.664**
HGNS without 3D-U-Net (only SGNN)	74.8	0.307	0.307	110.5	0.829	0.829
HGNS without SGNN (only 3D-U-Net)	94.6	0.361	0.361	142.6	0.725	0.724
HGNS with 1-step	**47.3**	0.500	0.500	122.8	1.144	1.144

Table 2 Ablation of our HGNS model

Rollout visualization on one trajectory, at a depth=10:



(a) HGNS rollout of water volume (barrel) for 20 steps (months)



(b) Ground-truth of water volume (barrel) for 20 steps

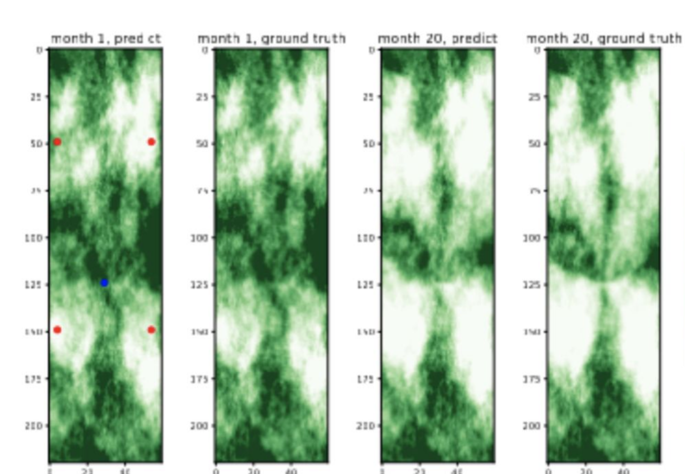


Fig. 4 Oil (barrel) prediction vs. ground truth

Speed comparison:
For rollout 20-step, 1.1 million cells time step:

- HGNS: 20.7s with an NVIDIA Quadro RTX 8000 GPU
- Solver: 46s-370s (varying depending on the number of wells) with 4 compute nodes, each with 2 CPUs Intel(R) Xeon(R) E5-2680 v3 2.50GHz

HGNS has 2 to 18-fold reduction in execution time.

For more information, see our paper "Learning Large-scale Subsurface Simulations with a Hybrid Graph Network Simulator" at
https://ai4earthscience.github.io/iclr-2022-workshop/