SPE · AAPG · SEG
CCUS
Carbon Capture, Utilization, and Storage

CCUS: 4443499

Benchmarking Advanced Graph Neural Networks for CO2 Plume Migration in Complex Geological Formations

Rodrigo S. Luna[1,2], Thiago H. N. Coelho[1,2], Luiz S. L. Neto[2,4], Roberto M. Velho[1,2], Adriano M. A. Cortes[1,2], Renato N. Elias[2,4], Alexandre G. Evsukoff[2,4], Fernando A. Rochinha[2,3], Herve Gross[5], Mauricio Araya-Polo*[5], Alvaro L. G. A. Coutinho[2,4], 1. Systems and Computer Engineering, COPPE/Federal University of Rio de Janeiro, 2. HPC Center, NACAD-COPPE/Federal University of Rio de Janeiro, 3. Mechanical Engineering, COPPE/Federal University of Rio de Janeiro, 4. Civil Engineering, COPPE/Federal University of Rio de Janeiro, 5. TotalEnergies EP R&T US.

Abstract

Accurate simulation of CO_2 migration in geological storage formations requires solving strongly coupled, nonlinear multiphase flow equations in heterogeneous porous media. While high-fidelity numerical simulators based on Darcy-scale models provide reliable predictions, their computational cost limits their use in repeated or real-time analyses. This paper presents a graph neural network (GNN) surrogate designed to approximate multiphase flow dynamics relevant to carbon capture and storage (CCS). The proposed formulation respects mesh connectivity and geometric structure, incorporates anisotropic message passing to emulate directional transport, and employs recurrent temporal integration to capture density-driven instabilities. The surrogate jointly predicts gas saturation and liquid-phase density and is validated on the SPE11a benchmark, which features sharp gas–water interfaces and convective fingering. A comparison of three training strategies, covering injection-only, post-injection, and combined regimes, shows that regime-specific models significantly improve long-horizon forecasting accuracy. The results indicate that the proposed approach can serve as an efficient complement to high-fidelity reservoir simulators for CCS scenario analysis.

Introduction

Predicting CO_2 plume evolution in subsurface formations requires solving coupled mass conservation equations for multiphase, multicomponent flow in porous media, combined with constitutive relations for relative permeability, capillary pressure, and equations of state. High-fidelity reservoir simulators based on finite-volume or finite-element discretizations provide reliable predictions but demand substantial computational resources, especially when fine resolutions are needed to capture sharp fronts or convective instabilities. To mitigate these costs, data-driven surrogate models have gained attention. Among them, graph-based methods are particularly appealing, as they naturally align with mesh-based numerical schemes, where control volumes map to nodes and flux interactions to edges (Pfaff et al., 2021), and can learn nonlinear transport phenomena directly from data without linearization or modal truncation.

This work presents a graph-based surrogate tailored to CO_2 storage problems (Ju et al., 2024), with emphasis on directional transport, injection-to-post-injection regime transitions, and density-driven flow. The SPE11a benchmark (Landa-Marbán and Sandve, 2025; Nordbotten et al., 2024; Rasmussen et al., 2021) serves as the test case, given its sharp gas–liquid interface and rapid onset of convective fingering.

Theory and Numerical Analogy

The computational mesh is mapped to a graph where each control volume corresponds to a node, and edges encode transmissibility-consistent neighborhood relationships. Node variables represent cell-centered quantities (gas saturation, liquid-phase density), while edge attributes capture geometric information analogous to face normals, distances, and interface areas, establishing a direct analogy between graph message passing and numerical flux exchange. Since multiphase transport is highly anisotropic, driven by permeability tensors, gravity, and density contrasts, the surrogate employs anisotropic message passing (Thürlemann and Riniker, 2023): neighboring influences are weighted by geometric and physical edge features, allowing preferential information propagation along dominant flow directions. Residual updates across message-passing layers parallel iterative flux evaluations and improve stability of the learned dynamics.

Time evolution is handled by a recurrent mechanism on latent node representations, acting as a learned time integrator that predicts state increments rather than absolute values, mirroring incremental time-stepping schemes and maintaining stability over long horizons. The model is validated on the SPE11a benchmark (Nordbotten et al., 2024), a laboratory-scale CO_2 injection problem featuring immiscible and partially miscible two-phase flow, strong density contrasts, and convective fingering during the post-injection phase, making it a stringent test for both numerical solvers and surrogates.

Results

The graph-based surrogate is trained on high-fidelity simulation trajectories and evaluated on a geological configuration unseen during training. The simulation spans 5 days and comprises two physically distinct phases: a 5-hour injection period followed by post-injection plume evolution. To account for these different dynamics, three models are considered: one trained exclusively on the injection phase, one on the post-injection phase, and one on the entire trajectory without regime distinction. Predicted gas saturation and liquid-phase density fields are compared against reference solutions across both flow regimes.

As shown in Figure 1, the surrogate captures the spatial extent of the CO_2 plume, the emergence of fingering patterns, and the gradual transition toward steady-state behavior. Quantitative assessment through RMSE ($\times 10^{-3}$) over increasing forecasting horizons shows that, for single-step predictions, all

three training strategies yield comparable accuracy, with gas saturation (S^g) errors below 1.2 and liquid-phase density ($\rho\ell$) errors below 1.0. At 10 steps, the injection-only model exhibits markedly higher errors (S^g: 15.22 ± 4.75; $\rho\ell$: 10.74 ± 2.86), whereas the post-injection (S^g: 4.80 ± 4.71; $\rho\ell$: 3.25 ± 2.67) and combined (S^g: 4.95 ± 4.62; $\rho\ell$: 3.43 ± 2.97) models remain substantially more accurate. At 50 steps, this gap widens further: the injection-only model reaches 85.35 ± 21.41 for S^g and 71.47 ± 10.34 for $\rho\ell$, while the post-injection and combined models stay in the 23–25 range for S^g and 16–17 for $\rho\ell$. These results confirm that regime-specific training significantly improves long-horizon forecasting, and that the combined model achieves accuracy close to the dedicated post-injection surrogate.

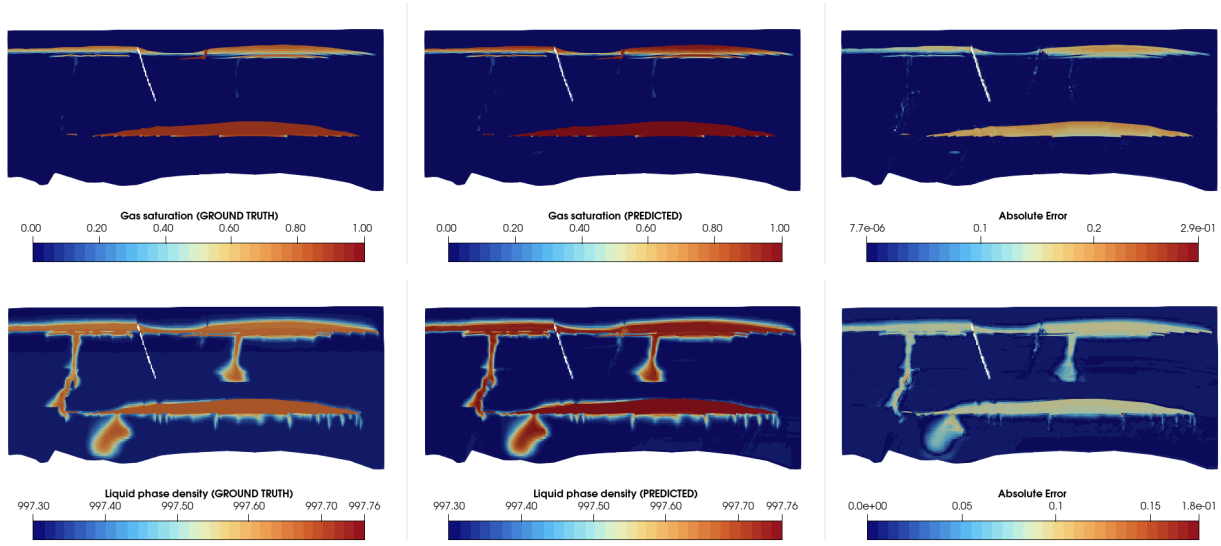

Figure 1. Reference porous media flow solution (left) versus graph-based surrogate prediction (middle) and absolute error (right) for gas saturation (top, dimensionless) and liquid-phase density (bottom, $kg \cdot m^{-3}$) at t = 23,400 s, shortly after the end of CO_2 injection. Inference starts at t = 22,500 s using the post-injection regime model.

Discussion

The results indicate that graph-based models can approximate the complex evolution of multiphase flow when trained on sufficiently rich datasets, serving as efficient surrogates for computationally expensive high-fidelity simulators. Several components contribute to the effectiveness of the proposed approach. The anisotropic message-passing mechanism is essential for reproducing transport-dominated phenomena, while latent temporal integration helps control error accumulation in transient simulations. Furthermore, the joint prediction of gas saturation and liquid-phase density within a single framework enables the surrogate to capture the coupling between phase distribution and density-driven convective mixing, which governs the post-injection dynamics of CO_2 storage.

The comparison across training regimes reveals that models trained on regime-specific data outperform those trained on the full trajectory, particularly over long forecasting horizons, suggesting that the distinct physical behaviors governing injection and post-injection phases benefit from dedicated surrogate representations. The combined model, while slightly less accurate than the dedicated post-injection surrogate, offers a practical compromise by providing reasonable predictions across both regimes without requiring explicit regime switching. Unlike traditional reduced-order models, the proposed approach does not rely on linear assumptions or modal bases, making it applicable to strongly nonlinear regimes; however, its accuracy remains sensitive to the representativeness of the training data, particularly across different flow regimes.

Conclusions

This paper presents a graph-based surrogate model for the simulation of CO_2 injection and storage in complex geological subsurface reservoirs. By aligning the graph formulation with mesh-based discretizations and incorporating anisotropic transport and recurrent time integration, the surrogate reproduces key features of multiphase flow observed in high-fidelity simulations of the SPE11a benchmark. The results indicate that such models can serve as efficient reduced-order approximations for CCS simulations and scenario analysis, complementing rather than replacing high-fidelity reservoir simulators. Since surrogate construction relies on rich data under supervised learning, future work will focus on tighter coupling with physical constraints, which may reduce the volume of training data required, extension to three-dimensional problems, systematic error control, and exploration of alternative architectures and training strategies.

Acknowledgements

We gratefully acknowledge the technical collaboration provided by TotalEnergies. This research was partially sponsored by TotalEnergies E&P Brazil under the 1% ANP obligation (GNN CO2 Project 24630-6). We also acknowledge that this investigation was financed in part by the Brazilian agency: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil - Finance code 001.

References

Ju, X., Hamon, F. P., Wen, G., Kanfar, R., Araya-Polo, M., & Tchelepi, H. A. 2024. Learning co2 plume migration in faulted reservoirs with graph neural networks. Computers Geosciences, 193, 105711. Retrieved from https://www.sciencedirect.com/science/article/pii/S0098300424001948 doi:https://doi.org/10.1016/j.cageo.2024.105711

Landa-Marbán, D., & Sandve, T. H. 2025. pyopmspe11: A python framework using opm flow for the spe11 benchmark project. Journal of Open-Source Software, 10(105), 7357. Retrieved from https://doi.org/10.21105/joss.07357 doi: 10.21105/joss.07357

Nordbotten, J. M., Ferno, M. A., Flemisch, B., Kovscek, A. R., & Lie, K.-A. 2024. The 11th society of petroleum engineers' comparative solution project: Problem definition. SPE Journal, 29(05), 2507–2524.

Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., & Battaglia, P. W. 2021. Learning mesh-based simulation with graph networks. In the International Conference on Learning Representations (ICLR). Retrieved from https://arxiv.org/abs/2010.03409

Rasmussen, A. F., Sandve, T. H., Bao, K., Lauser, A., Hove, J., Skaflestad, B., et al. 2021. The open porous media flow reservoir simulator. Computers Mathematics with Applications, 81, 159-185. Retrieved from https://www.sciencedirect.com/science/article/pii/S0898122120302182 (Development and Application of Open-source Software for Problems with Numerical PDEs) doi:https://doi.org/10.1016/j.camwa.2020.05.014

Thürlemann, M., & Riniker, S. 2023. Anisotropic message passing: Graph neural networks with directional and long-range interactions. In The eleventh international conference on learning representations.