

CCUS: 4413231

Integrated Reduced-Order Modeling and Reinforcement Learning for Intelligent Geological Carbon Storage Management: A Real-Time Optimization Framework for Complex 3D Heterogeneous Reservoirs

T. A. Yehia^{*1}, Y. Falola², I. Eloghosa³, G. G. Nair², J. Toms², D. N. Meehan¹, 1. Texas A&M University, 2. Louisiana State University, 3. Halliburton.

Copyright 2026, Carbon Capture, Utilization, and Storage conference (CCUS) DOI 10.15530/ccus-2026-4413231

This paper was prepared for presentation at the Carbon Capture, Utilization, and Storage conference held in The Woodlands, TX, 30 March – 01 April.

The CCUS Technical Program Committee accepted this presentation on the basis of information contained in an abstract submitted by the author(s). The contents of this paper have not been reviewed by CCUS and CCUS does not warrant the accuracy, reliability, or timeliness of any information herein. All information is the responsibility of, and, is subject to corrections by the author(s). Any person or entity that relies on any information obtained from this paper does so at their own risk. The information herein does not necessarily reflect any position of CCUS. Any reproduction, distribution, or storage of any part of this paper by anyone other than the author without the written consent of CCUS is prohibited.

Abstract

This study couples 3D Embed-to-Control-and-Observe (E2CO) reduced-order modeling with reinforcement learning (RL) to optimize geological CO₂ storage operations, maximizing injection efficiency while ensuring well-level safety and regulatory compliance. The RL agent uses a reward function promoting higher CO₂ injection and reduced brine handling, with strict penalties for plume migration near producer wells or excessive reservoir pressures. A 30-year storage project was simulated using the heterogeneous SACROC reservoir model with three injectors and three producers across three geological realizations (1–5 MtCO₂/year). The E2CO architecture achieved ~10,000× acceleration over full-physics models by embedding flow dynamics in latent space, enabling training over 1,000 RL episodes within minutes. The RL framework successfully learned adaptive strategies exploiting geological heterogeneity—increasing injection in high-capacity regions while reducing rates near critical thresholds. Hard penalties ensured operational safety, and learned policies generalized effectively across geological realizations. Compared to static approaches, RL-optimized strategies achieved higher cumulative injection, improved containment security, and more stable pressure management while supporting continuous optimization as monitoring data became available. This work integrates RL with E2CO for CO₂ storage optimization, enabling real-time, adaptive reservoir management that generalizes across uncertain geological realizations.

Introduction

Surrogate modeling has attracted significant attention across many scientific and engineering disciplines, particularly in subsurface reservoir engineering, where high-fidelity numerical simulators are computationally intensive (Nanga et al. 2021). Rapid surrogate predictions enable large-scale uncertainty

quantification, real-time monitoring, and iterative optimization workflows that would otherwise be infeasible if every evaluation required running a full-physics simulator. As carbon storage management problems increasingly demand repeated forward simulations, the development of accurate and efficient reduced-order (ROMs) and machine learning models has become a critical research focus (Falola et al. 2023; 2024; 2024; 2025; 2026; Abdulkareem et al. 2025).

Traditional ROM techniques like proper orthogonal decomposition (POD) offer speedups but rely on intrusive linearization and struggle with large control perturbations or nonlinear regimes (Gratton and Willcox 2004; He and Durlofsky 2013). Deep learning surrogates provide greater flexibility, but early versions lacked control-aware, long-horizon forecasting mechanisms. The Embed-to-Control (E2C) framework introduced control-compatible deep surrogates by learning low-dimensional latent representations with locally linear dynamics, enabling classical optimal control in latent space (Watter et al. 2015). However, initial formulations weren't designed for subsurface flow systems with physics-driven dynamics and well-based controls (Jin et al. 2020). Reservoir-adapted E2C extended the framework to multiphase flow, achieving orders-of-magnitude speedups with stable multi-step rollouts, but required explicit well models like Peaceman formulations, introducing errors and overhead. The Embed-to-Control-and-Observe (E2CO) framework addressed this by augmenting the latent model with a learned observation operator that directly predicts well outputs from latent transitions (Coutinho et al. 2021). E2CO has been extended to 3D heterogeneous reservoirs, large-scale geological models, and ensemble formulations accounting for geological uncertainty (Atadeger et al. 2022; 2023; Aghayev et al. 2024).

Fast, differentiable E2CO surrogates enable integration with gradient-based constrained optimization and deep reinforcement learning (DRL). Coupling E2CO with stochastic gradient-based sequential quadratic programming (SQP) enables robust life-cycle optimization under nonlinear constraints (Nguyen et al. 2024), while E2CO-based environments accelerate RL by allowing agents to explore policies in the learned latent space (Chen et al. 2024; Wang et al. 2025; Adeyemi and Onur 2025). This work extends 3D-E2CO to geological carbon storage by integrating multi-realization latent surrogates with deep reinforcement learning for robust optimization of CO₂ injection and pressure management under geological uncertainty in the SACROC unit reservoir.

Dataset and Methodology

The study uses the SACROC unit, a mature carbonate reservoir in the Permian Basin, West Texas, repurposed for CO₂ geological storage (Jia and McPherson 2018). The domain is a 34×16×25 grid (13,600 cells), spanning ~8,400×4,000 m in area with a thickness of 27-89m across 25 layers. The configuration includes three CO₂ injectors (I1, I2, I4) and three producers (P1, P2, P3), all fully perforated. Three geological realizations represent subsurface uncertainty with distinct porosity and permeability fields. Extreme negative values replace inactive cells during preprocessing with masking to exclude them from reconstruction loss (Le and Tao 2024). A commercial compositional simulator generates the training dataset. For each realization, well controls, producer bottom-hole pressures, and CO₂ injection rates (0.5-5 megaton/year), are varied using Latin Hypercube sampling over 30 years with annual control intervals, yielding 1,000 simulation cases executed in ~72 hours on a multi-core Intel Xeon workstation with 768 GB RAM. At each time step, 4 spatial state fields (reservoir pressure, gas saturation, porosity, permeability) and 9 well-level time-series observations (three injector BHPs, three gas production rates, three water production rates) are recorded. Spatial properties are normalized, and data is partitioned into 75% training and 25% testing sets.

The ROM follows the E2CO paradigm with three components: an encoder using 3D convolutional layers to compress reservoir states (34×16×25 grid, 4 channels) into a 128-dimensional latent vector; a decoder expanding the latent vector back to original resolution via transposed convolutions; and a linear transition model predicting next latent state and 9 well-level observations (3 injector BHPs, 3 gas/water production rates) from current latent state and 6 well controls (3 producer BHPs, 3 CO₂ injection rates). The trained 3D-E2CO surrogate serves as the environment for a model-free RL agent learning optimal well-control policy over the 30-year CO₂ injection horizon, with the reservoir control problem cast as a Markov

Decision Process where the agent observes latent reservoir state at each annual decision step $t = 1 \dots 30$, the agent observes the latent reservoir state z_t (a 128-dimensional vector produced by the E2CO encoder), selects a 6-dimensional action a_t comprising three producer bottom-hole pressures and three CO₂ injection rates (all normalized to $[0, 1]$ and mapped to physical ranges by the environment), and receives an economic reward. The surrogate then advances the latent state via the learned transition model to produce z_{t+1} and the corresponding 9-dimensional well observations y_{t+1} . At each episode reset, an initial state z_0 is randomly sampled from the available geological realizations, exposing the agent to subsurface uncertainty during training. The per-step reward captures the net economic value of CO₂ geological storage:

$$r_t = \frac{1}{S} \left(c_{inj} \cdot \eta \cdot Q_{inj,t} - c_w \cdot \frac{Q_{w,t}}{\beta} - c_{gp} \cdot \eta \cdot Q_{gp,t} \right)$$

Where $Q_{inj,t}$ is the total CO₂ injection rate, $Q_{w,t}$ is the total water production rate, $Q_{gp,t}$ is the total gas production rate, eta is the gas unit conversion factor (ft^3/day to tons/day), beta = $5.615 ft^3/bbl$ is the volumetric conversion, $c_{inj} = \$30/ton$ is the net CO₂ storage credit (revenue minus operating cost), $c_w = \$5/bbl$ is the water production penalty, $c_{gp} = \$50/ton$ is the gas breakthrough penalty, and $S = 10^{-4}$ S is a normalization scale factor.

This integrated ROM-RL framework enables the optimization of long-horizon well control strategies in seconds rather than hours, while explicitly accounting for the geological uncertainties that dominate subsurface decision-making.

Results and Discussion

The SAC agent demonstrated strong convergence over 1,000 episodes (Fig.1a) with rapid learning within the first 200 episodes, reaching stability around episode 326, with sustained high-reward plateau through episode 1,000, confirming policy robustness against geological uncertainty. Lifecycle analysis (Fig.1b) shows rapid recovery from initial capital expenditure with monotonically increasing cumulative cashflow through the 35-year project life, while NPV sensitivity analysis (Fig.1c) reveals profitability at discount rates up to 10%, with higher rates yielding negative NPVs—characteristic of capital-intensive, long-horizon carbon storage projects. The RL agent manages the 30-year operational horizon (Fig.2) using dynamic BHP controls for three production wells (P1–P3) within a 1,150–1,350 psi range, resulting in production profiles where declining water production and rising gas rates signal progressive CO₂ breakthrough with varied timings reflecting geological heterogeneity. To maximize storage, the agent sustains gas injection rates near upper bounds for wells I1–I3 while proactively introducing periodic sharp injection reductions—specifically in I2 and I3—to manage pressure buildup and mitigate early breakthrough, with injector BHP observations revealing declining trends interspersed with transient spikes correlating with strategic rate adjustments, demonstrating responsive control logic driven by six control actions refined by nine observation channels.

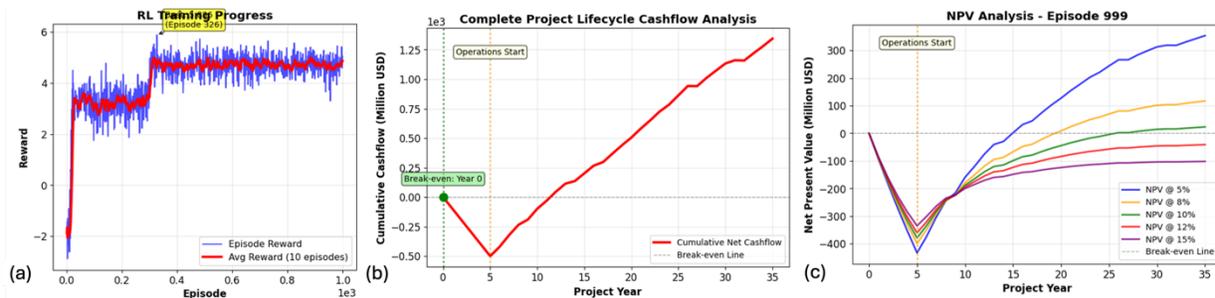


Figure 1. Training performance and economic analysis of the SAC-based well control optimization across 1,000 geological realizations: (a) episodic and 10-episode moving average reward showing convergence by episode 326; (b) cumulative project lifecycle cashflow over 35 years under the optimized policy; (c) NPV sensitivity to discount rate (5-15%) at the final training episode.

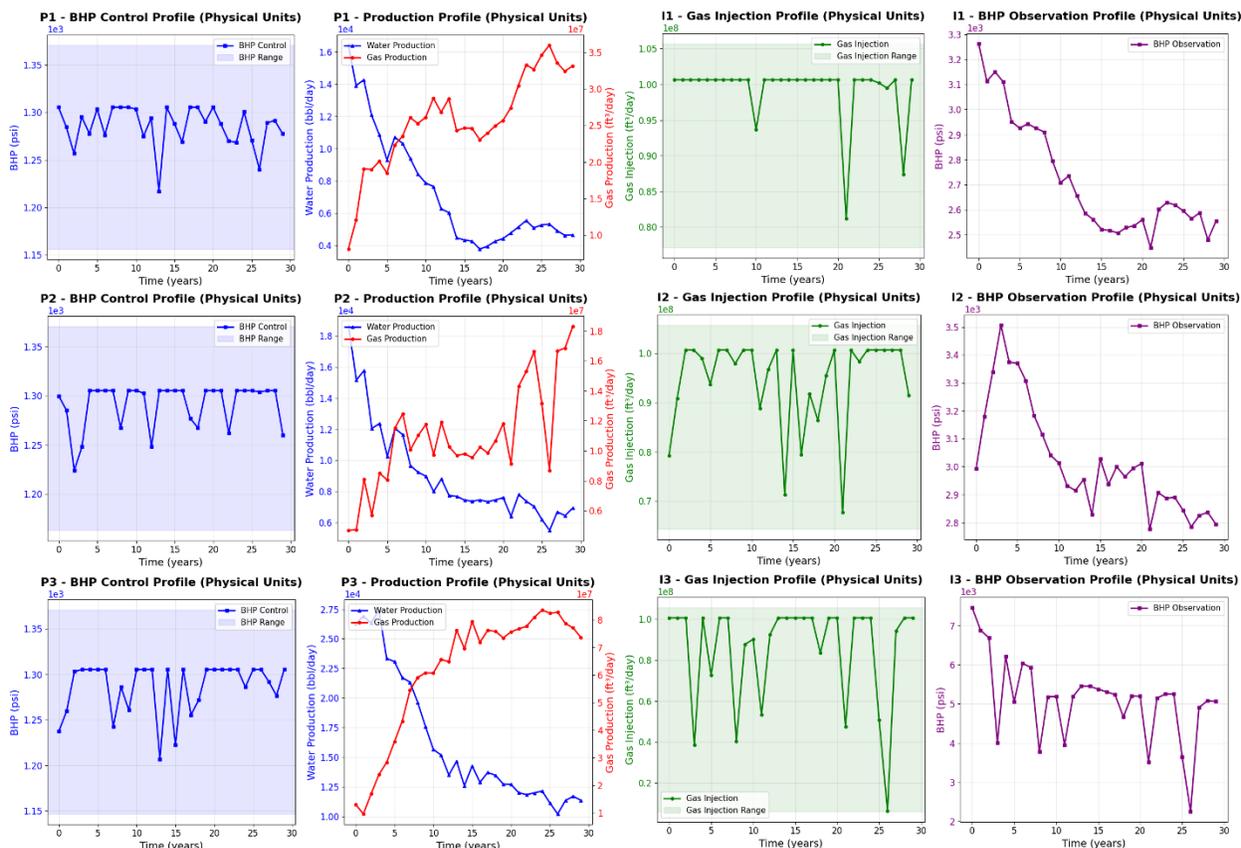


Figure 2. RL-optimized well control actions and observed reservoir responses over a 30-year horizon for a representative geological realization: producer BHP controls and associated water/gas production profiles (left), and injector gas injection rate controls with resulting BHP observations (right). Shaded regions denote the permissible operational ranges for each control variable.

Conclusions

The integration of the 3D surrogate framework with RL establishes an efficient, real-time optimization engine for geological carbon storage, as demonstrated on the SACROC model unit. The SAC agent’s rapid convergence by episode 326 and sustained high-reward plateau through 1,000 episodes confirm a policy that is robust against subsurface uncertainty and generalizes effectively across diverse geological realizations. By autonomously adapting well-level controls to exploit reservoir heterogeneity and mitigate early CO₂ breakthrough, the framework achieves superior sequestration efficiency while maintaining

operational safety. Future research should explore hybrid optimization approaches, where RL is coupled with other optimizers to further refine control policies over the surrogate model. Subsequent work must focus on the rigorous comparison and validation of these hybrid strategies against standalone RL to ensure maximum reliability in large-scale sequestration projects.

References

- Abdulkareem, Usman, Ahmed Adeyemi, and Mustafa Onur. 2025. "History Matching Geological Models Using an Embed-To-Control Observe Deep-Learning Reservoir Surrogate." <https://dx.doi.org/10.2118/228029-MS>.
- Adeyemi, Ahmed, and Mustafa Onur. 2025. "Accelerated Deep Reinforcement Learning in Subsurface Reservoir Production Optimization." *SPE Journal* 30 (11): 6603–21. <https://doi.org/10.2118/230310-PA>.
- Aghayev, Zahir, Dimitrios Voulanas, Eduardo Gildin, and Burcu Beykal. 2024. "Enhancing Oil Recovery Efficiency with Constrained Reinforcement Learning for Waterflooding Optimization." Paper presented at 2024 AIChE Annual Meeting, October 28. <https://aiche.confex.com/aiche/2024/meetingapp.cgi/Paper/689002>.
- Atadeger, A., S. Sheth, G. Vera, R. Banerjee, and M. Onur. 2022. "Deep Learning-Based Proxy Models to Simulate Subsurface Flow of Three-Dimensional Reservoir Systems." *ECMOR 2022*, 1–32. <https://doi.org/10.3997/2214-4609.202244049>.
- Atadeger, Aykut, Mustafa Onur, Soham Sheth, and Raj Banerjee. 2023. "Deep Learning-Based and Kernel-Based Proxy Models for Nonlinearly Constrained Life-Cycle Production Optimization." <https://dx.doi.org/10.2118/212690-MS>.
- Chen, Jungang, Eduardo Gildin, and Georgy Kompantsev. 2024. "Optimization of Pressure Management Strategies for Geological Co2 Sequestration Using Surrogate Model-Based Reinforcement Learning." Preprint, SSRN. <https://doi.org/10.2139/ssrn.4806124>.
- Coutinho, Emilio Jose Rocha, Marcelo Dall'Aqua, and Eduardo Gildin. 2021. "Physics-Aware Deep-Learning-Based Proxy Reservoir Simulation Model Equipped With State and Well Output Prediction." *Frontiers in Applied Mathematics and Statistics* 7 (September): 651178. <https://doi.org/10.3389/fams.2021.651178>.
- Falola Yusuf, Siddharth Misra, and Andres Nunez. 2023. "Rapid high-fidelity forecasting for geological carbon storage using neural operator and transfer learning", Abu Dhabi International Petroleum Exhibition and Conference. <https://doi.org/10.2118/216135-MS>.
- Falola Yusuf, Siddharth Misra, and Andres Nunez. 2024. "Integrating Neural Operators and Transfer Learning for Efficient Carbon Storage Forecasting", Abu Dhabi International Petroleum Exhibition and Conference. <https://doi.org/10.2118/222406-MS>.
- Falola Yusuf, Pradyumna Singh Rathore, Geetha Gopakumar Nair, and Julianna Toms. 2024. "Application of Neural Operator Technique for Rapid Forecast of CO2 Pressure and Saturation Distribution." <https://dx.doi.org/10.4043/35230-MS>.

Yusuf Falola, Siddharth Misra, and Andres Nunez. 2025. “Geological and Engineering Insights from Transfer Learning with Fourier Neural Operators: A Case Study of CO₂ Storage Forecasting in Disparate Saline Aquifers.” *Applied Soft Computing* (178), <https://doi.org/10.1016/j.asoc.2025.113272>.

Yusuf Falola, Siddharth Misra, and Andres Nunez. 2026. “Leveraging transfer learning for reliable CO₂ storage forecasting across diverse operational conditions.” *Fuel* (403), <https://doi.org/10.1016/j.fuel.2025.135990>.

Gratton, David, and Karen E. Willcox. 2004. Reduced-Order, Trajectory Piecewise-Linear Models for Nonlinear Computational Fluid Dynamics. January. <https://dspace.mit.edu/handle/1721.1/3892>.

Jia, Wei, and Brian McPherson. 2018. A CMG-GEM Model for Simulating CO₂-EOR at SACROC Unit. 1 (January). <https://doi.org/10.17632/cxftbmzzkx.1>.

Jin, Zhaoyang Larry, Yimin Liu, and Louis J. Durlofsky. 2020. “Deep-Learning-Based Surrogate Model for Reservoir Simulation with Time-Varying Well Controls.” *Journal of Petroleum Science and Engineering* 192 (September): 107273. <https://doi.org/10.1016/j.petrol.2020.107273>.

Le, Hieu, and Jian Tao. 2024. “Hierarchical Autoencoder-Based Lossy Compression for Large-Scale High-Resolution Scientific Data.” *Computing&AI Connect* 1 (1): 1. <https://doi.org/10.69709/CAIC.2024.193132>.

Nanga, Salifu, Ahmed Tijani Bawah, Benjamin Ansah Acquaye, et al. 2021. “Review of Dimension Reduction Methods.” *Journal of Data Analysis and Information Processing* 09 (03): 189–231. <https://doi.org/10.4236/jdaip.2021.93013>.

Nguyen, Quang Minh, Mustafa Onur, and Faruk Omer Alpak. 2024. “Embed-to-Control-Based Deep-Learning Surrogate for Robust Nonlinearly Constrained Life-Cycle Production Optimization: A Realistic Deepwater Application.” <https://dx.doi.org/10.2118/220783-MS>.

Wang, Zhongzheng, Yuntian Chen, Guodong Chen, and Dongxiao Zhang. 2025. “Optimization of Geological Carbon Storage Operations with Multimodal Latent Dynamic Model and Deep Reinforcement Learning.” *Geoenergy Science and Engineering* 244 (January): 213407. <https://doi.org/10.1016/j.geoen.2024.213407>.

Watter, Manuel, Jost Springenberg, Joschka Boedecker, and Martin Riedmiller. 2015. “Embed to Control: A Locally Linear Latent Dynamics Model for Control from Raw Images.” In *Advances in Neural Information Processing Systems*, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, vol. 28. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2015/file/a1afc58c6ca9540d057299ec3016d726-Paper.pdf.