geoconvention
Calgary • Canada • May 11-13 2026

From Signal to Strategy: Geophysical Control of Water-Alternating-Gas Injection Strategies and Carbon Sequestration via Deep Reinforcement Learning (DRL)

EvelynLove Fosu-Duah[1], Luis E. Zerpa[2], Kyubo Noh[1], Andrei Swidinsky[1]
University of Toronto[1], Colorado School of Mines[2]

Introduction

CO_2 Water-Alternating-Gas (WAG) enhances oil recovery by improving sweep efficiency, reducing viscous fingering and gas channelling while enabling long-term CO_2 storage in depleted reservoirs. However, optimizing CO_2-WAG operations is challenging due to nonlinear, time-dependent reservoir behaviour and competing objectives between oil recovery and carbon storage. Conventional, heuristic-based designs are often inefficient and costly, particularly when short-term production trade-offs are required. This study applies deep reinforcement learning (DRL) to enable adaptive, sequential optimization of CO_2-WAG injection strategies under evolving subsurface conditions.

Methodology

We formulate CO_2-WAG optimization as a Markov Decision Process in which a DRL agent interacts with a deterministic reservoir–geophysical simulation to maximize cumulative Net Present Value (NPV). The reservoir is modelled using the Open Porous Media Flow Simulator. It represents a homogeneous, anisotropic oil reservoir discretized on a 30 × 30 × 5 grid, with an injector and producer located at opposite corners. Each decision step consists of a reservoir flow simulation followed by forward gravity modelling.

The agent's state is defined by surface-time-lapse gravity observations that capture subsurface density changes associated with water, oil, and CO_2 migration and redistribution. Two observation configurations are considered: a single central gravimeter and a 60 × 60 grid surface gravity array. Temporal information is incorporated by using gravity measurements from consecutive time steps. The reward function is defined as discounted NPV, accounting for oil revenue and water and CO_2 injection and production costs. The action space comprises 11 discrete injection strategies, each corresponding to a combination of water and CO_2 injection rates.

Two RL approaches are evaluated: tabular Q-Learning (QL) and Double Deep Q-Networks (DDQN). QL is tested using coarse- and fine-resolution discretization of gravity observations to assess the impact of state granularity. DDQN addresses the resulting scalability limitations by leveraging function approximation with artificial neural networks (ANNs) and convolutional neural networks (CNNs), enabling learning from high-dimensional spatial gravity data. Agents are trained for 2,000 episodes with bi-annual decision intervals, allowing systematic comparison of learning performance and control effectiveness under partial observability.

Results

Q-learning (QL) and Double Deep Q-Network (DDQN) methods produced distinct injection policies, highlighting the influence of state representation and algorithm design (Figure 1). Finer state discretization improved control fidelity by capturing subtle variations in the reservoir. However, it increased training time and computational cost due to sparse state visitation, whereas coarser discretization accelerated convergence at the expense of policy optimality (Sutton & Barto, 2018).

When benchmarked against industrial schedules, water flooding, CO_2 flooding, and biannual CO_2-WAG, QL outperformed standalone water and CO_2 flooding but underperformed the industrial WAG strategy (Figure 1). High-resolution Q-learning consistently outperformed coarse-state discretization, demonstrating that finer state representations more effectively capture the underlying reservoir dynamics, in agreement with prior findings (Fosu-Duah et al., 2025).

To address QL limitations, DDQN was implemented with both ANN and CNN function approximators. ANN-based DDQN underperformed QL, consistent with previous observations that increased model complexity does not necessarily improve performance in low-dimensional state spaces (van Hasselt et al., 2016; Figure 1). In contrast, CNN-based DDQN achieved the highest cumulative NPV by exploiting spatial correlations in time-lapse gravity data and reducing value-estimation bias (Figure 1).

A comparison of optimized and industrial WAG policies shows that the CNN-DDQN strategy increases cumulative CO_2 storage (Figure 2) by optimizing injection timing and reducing alternation frequency. Time-lapse gravity anomalies over 20 years further reveal sustained negative anomalies under the CNN-DDQN policy, indicating prolonged CO_2 residence and enhanced trapping relative to the industrial schedule (Figure 3), consistent with established CO_2 trapping mechanisms (Juanes et al., 2006).

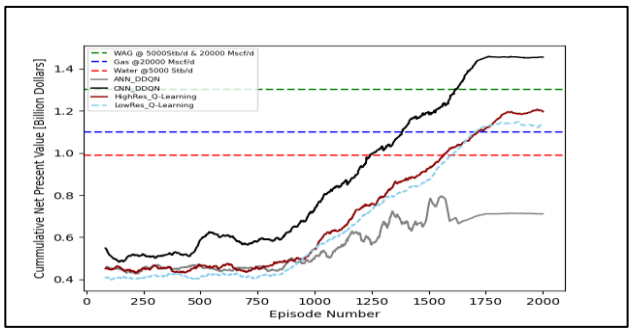


Figure 1: Industrial vs RL performance

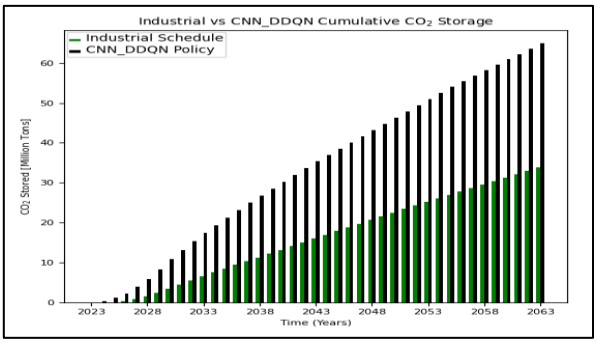


Figure 2: Industrial vs DDQN Cumulative CO_2 storage

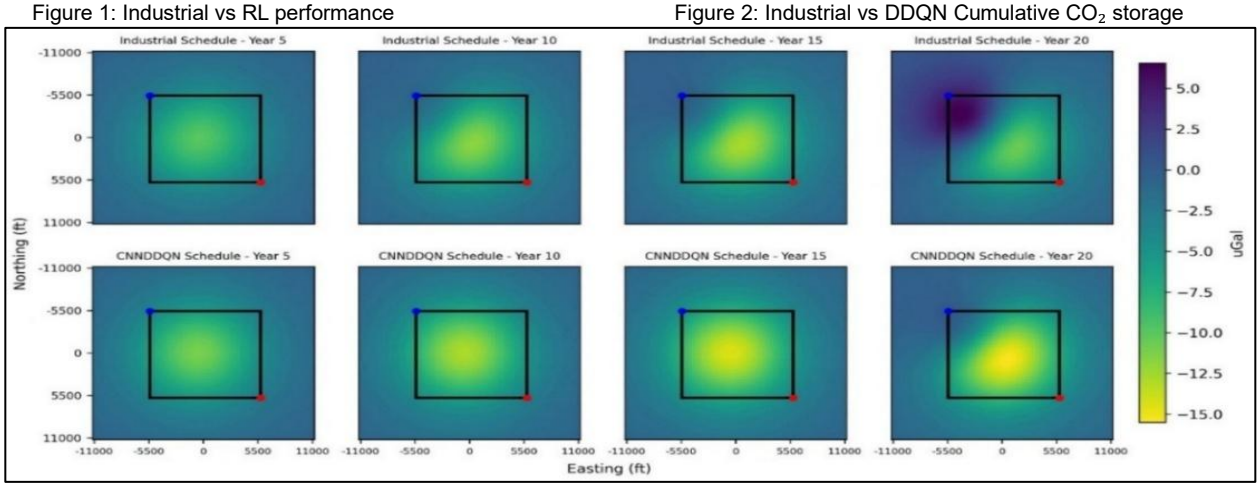


Figure 3: Industrial vs DDQN Time-lapse gravity anomaly evolution

Contribution

RL using time-lapse gravity data optimizes WAG injection, thereby improving CO_2 storage and economic performance relative to industrial schedules.

References

Fosu-Duah, E., Noh, K., Zerpa, L. E., and A. Swidinsky. "Water-Alternating-Gas and CO_2 Storage Optimization Using Time-Lapse Geophysical Monitoring and Deep Reinforcement Learning." Paper presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, USA, October 2025. doi: https://doi.org/10.2118/228052-MS

Juanes, R., E. J. Spiteri, F. M. Orr, & M. J. Blunt (2006). Impact of relative permeability hysteresis on geological CO_2 storage: Water Resources Research, 42(12), W12418, doi: https://doi.org/10.1029/2005WR004806.

Sutton, R. S., & A. G. Barto (2018). Reinforcement learning: An introduction (2nd ed.): MIT Press.

Van Hasselt, H., A. Guez, & D. Silver (2016). Deep reinforcement learning with double Q-learning: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16), 2094–2100.