Learning Hydrogen Flow Behavior in Heterogeneous Saline Aquifers under Cyclic Operation with Physics-Constrained CNN-RNN Framework

Zhen Xu[1,2], Amin Misaghi Bonabi[3], Mengjie Zhao[3], Marc Gerritsma[4], Hadi Hajibeygi[5], Juan Alcalde[6], and Victor Vilarrasa[1]

[1]IMEDEA UIB-CSIC, Esporles, Spain [2]UPC-BarcelonaTech, Barcelona, Spain [3]TU Delft, Applied Mathematics, Delft, NL [4]TU Delft, Aerospace Engineering, Delft, NL [5]TU Delft, Civil Engineering, Delft, NL [6]GEO3BCN-CSIC, Barcelona, Spain

Motivation and Objectives

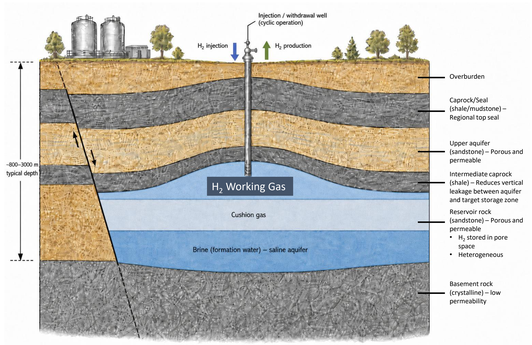

*Fig. 1. Schematic of underground hydrogen storage in a saline aquifer: cyclic H_2 injection/withdrawal beneath a sealing caprock, with structural trapping at the anticline crest and a normal sealing fault crossing the reservoir. **

Challenges:

- **Geological uncertainty challenges safe cyclic operation.** Caprock integrity, fault transmissivity, and reservoir heterogeneity govern H_2 containment, plume migration, and recoverability, yet are constrained only by sparse data. Hydrogen dissolution and residual trapping decrease the storage efficiency. Sufficient aquifer support, etc.
- **Conventional simulation is too slow for assessment.** Multiphase simulations with hysteresis, capillary, gravity and compositional effects must be rerun for every operational condition and realization, which are too expensive to account for real-time monitoring data.
- **A deep learning surrogate enables continual, dynamic forecasting.** A trained model predicts the new state of the system $[P, S]_{t+1}$ based on its past states at a fraction of the simulator cost and rolls out autoregressively over many future cycles.

Methodology

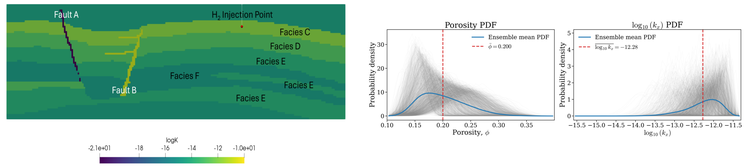

Fig. 2. (Left) 2D heterogeneous reservoir cross-section with six facies and two faults (A: sealing, B: conductive) [1]. (Right) Stochastic porosity and $\log_{10}(k_x)$ PDFs across the ensemble of realizations.

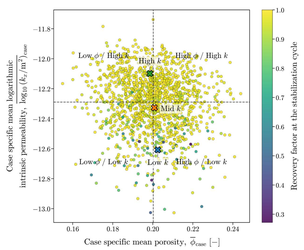

Fig. 3. Recovery factor (RF) at the stabilization cycle vs. rock properties across the realization ensemble. Each point is one stochastic realization; the spread reveals the sensitivity of H_2 recoverability to facies-scale heterogeneity.

Governing equations

$$\sum_{\alpha=w,n}\Big[\underbrace{\frac{\partial}{\partial t}(\phi S_\alpha \rho_\alpha x_{c,\alpha})}_{\text{accumulation}}+\underbrace{\nabla\cdot(\boldsymbol{u}_\alpha \rho_\alpha x_{c,\alpha})}_{\text{advection}}+\underbrace{\nabla\cdot(\rho_\alpha \dot{j}_{c,\alpha})}_{\text{diffusion}}+\underbrace{q_\alpha x_{c,\alpha}}_{\text{source/sink}}\Big]=0$$

$$\boldsymbol{u}_\alpha=-\frac{k_{r,\alpha}\mathbf{K}}{\mu_\alpha}(\nabla p_\alpha-\rho_\alpha g\nabla h) \qquad \dot{j}_{c,\alpha}=-\phi S_\alpha D_{c,\alpha}\nabla x_{c,\alpha}$$

1. **Synthetic geometry built on the SPE-11B benchmark.** A 2D anticline cross-section with six facies, a sealing fault (A) on the flank, and a conductive fault (B) near the axis, with H_2 injected into Facies C at the crest.
2. **H_2-brine flow simulated with DARSim two-phase module.** DARSim solves the component mass balance for wetting (brine) and non-wetting (H_2) phases with multiphase Darcy velocities and Fickian diffusion, providing the ground-truth [P, S] fields for surrogate training.
3. **Per-facies stochastic sampling of φ and k.** Porosity and log permeability are drawn independently per facies from prescribed distributions and re-targeted to the facies mean, so the realization suite spans wide intra-facies variability without bias.
4. **Recovery Factor (RF) is controlled mainly by permeability.** RF separates clearly along $\log_{10}(k_x)$ (high-k cases approach unity, low-k cases drop below 0.4), while the same porosity range produces only weak RF gradients at fixed k.

Result

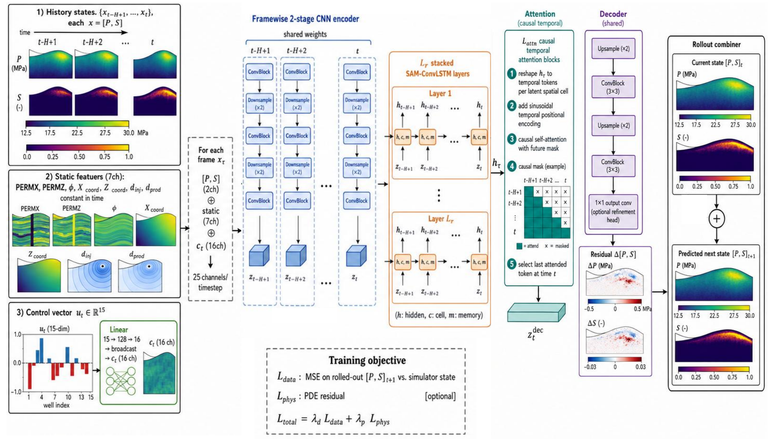

*Fig. 3. SAM-ConvLSTM rollout surrogate. (1) Inputs: history states $\{x_{t-H+1},\ldots,x_t\}$ = [P, S], 7 static features (PERMX, PERMZ, φ, X/Z coords, d_{inj}, d_{prod}) and a 15-d well control vector u_t embedded into a 16-channel field c_t; concatenated to 25 channels per timestep. (2) A weight-shared 2-stage CNN encodes each frame to a ¼-resolution latent. (3) L_r stacked SAM-ConvLSTM layers propagate spatio-temporal hidden state. (4) L_{attn} causal-attention blocks add sinusoidal time encoding and apply masked self-attention; the last attended token at time t feeds the decoder. (5) The decoder upsamples back to native resolution and predicts a residual Δ[P, S], which a rollout combiner adds to the current state to obtain the next state $[P, S]_{t+1}$. Training minimizes $L_{total} = \lambda_d L_{data} + \lambda_p L_{phys}$, with the PDE-residual term L_{phys} added in the physics-informed variant. **

Data-driven model

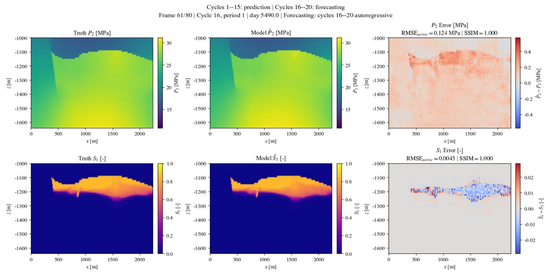

Fig. 4. Data-driven SAM-ConvLSTM forecast at cycle 16, day 5490 (autoregressive). Truth vs. model fields for P_2 and S_1 with error maps. RMSE_active = 0.124 MPa (P_2), 0.0045 (S_1); SSIM = 1.000.

Physics-informed model

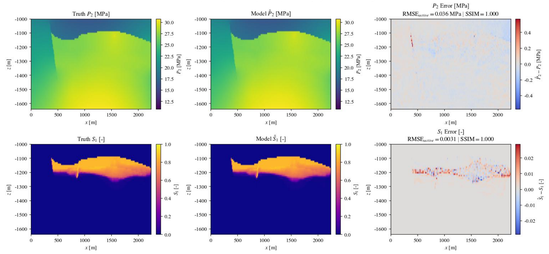

Fig. 5. Physics-informed SAM-ConvLSTM forecast at the same cycle/day. Physics loss reduces pressure forecast errors by ~3–4× comparing to pure data-driven: RMSE_active = 0.036 MPa (P_2), 0.0031 (S_1); SSIM = 1.000.

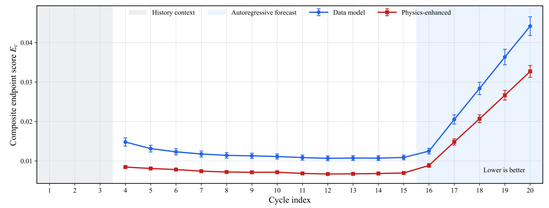

Fig. 6. Composite endpoint score E_c per cycle (mean ± std). Cycles 1–3 are history context; cycles 4–15 are one-step prediction; cycles 16–20 are autoregressive forecast. The physics-enhanced model maintains lower error throughout, with the gap widening as the rollout extends. Lower is better.

Four endpoint times per cycle: injection end, rest-1 end, production end, rest-2 end; $\mathcal{T}=\{1,2,3,4\}$ and $\tau\in\mathcal{T}$.

$u_{c,\tau,k}(x)$ is the reference field value and $\hat{u}_{c,\tau,k}(x)$ is the prediction at cell x, cycle c, endpoint τ.

k selects the physical field being scored; here $k\in\{P_2, S_1\}$, with S_1 the saturation field used for the plume mask.

$$E_c=\frac{1}{4}\left(E^{\text{global}}_{P_2,c}+E^{\text{global}}_{S_1,c}+E^{\text{plume}}_{P_2,c}+E^{\text{plume}}_{S_1,c}\right)$$

$$E^{\text{global}}_{k,c}=\frac{1}{R_k}\sqrt{\frac{1}{4|\Omega_a|}\sum_{\tau\in\mathcal{T}}\sum_{x\in\Omega_a}\left(\hat{u}_{c,\tau,k}(x)-u_{c,\tau,k}(x)\right)^2}$$

$$E^{\text{plume}}_{k,c}=\frac{1}{R_k}\sqrt{\frac{\sum_{\tau\in\mathcal{T}}\sum_{x\in\Omega_a}M_{c,\tau}(x)\left(\hat{u}_{c,\tau,k}(x)-u_{c,\tau,k}(x)\right)^2}{\sum_{\tau\in\mathcal{T}}\sum_{x\in\Omega_a}M_{c,\tau}(x)}}$$

$M_{c,\tau}(x)=1\{S_{1,c,\tau}(x)>0.05\}$, $R_k=P_{95}(u_k)-P_5(u_k)$, $\Omega_a=$ active cells.

Note: error norms are evaluated on the normalized state variables (pressure, saturation). Active cells are cells with a non-zero gas saturation change history.

Applications

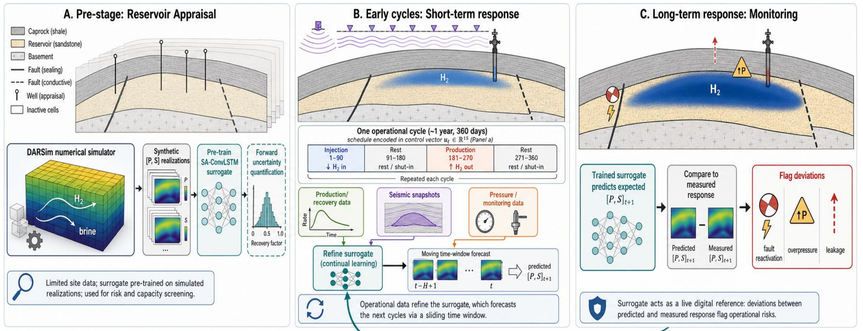

*Fig. 7. Workflow for developing and deploying a physics-constrained SA-ConvLSTM surrogate for cyclic underground hydrogen storage, from reservoir appraisal and early-cycle data assimilation to long-term monitoring and risk flagging. **

References

[1] Misaghi Bonabi, A.; van Rooijen, W.; Al Kobaisi, M.; Vuik, C.; Hajibeygi, H. Comparative analysis of carbon dioxide and hydrogen plume migration in aquifers inspired by the FluidFlower benchmark study. Int. J. Hydrogen Energy 2025, 135, 56–68.
[2] Zhao, M.; Wang, Y.; Gerritsma, M.; Hajibeygi, H. A physics-constraint neural network for CO2 storage in deep saline aquifers during injection and post-injection periods. Advances in Water Resources 2024, 193, 104837.

Declaration

[*] AI-assisted figure generated with OpenAI Image 2; reviewed by the author.

Acknowledgement

This work has received funding from the European Union's Horizon Europe Research and Innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101073271.

▶ Scan for animations
Cyclic H_2 injection & withdrawal forecasts