
ENHANCING ROBUSTNESS OF DIGITAL SHADOW FOR CO₂ STORAGE MONITORING WITH AUGMENTED ROCK PHYSICS MODELING

A PREPRINT

Abhinav Prakash Gahlot

Georgia Institute of Technology

Felix J. Herrmann

Georgia Institute of Technology

January 24, 2025

1 Introduction

In order to achieve climate targets, the IPCC emphasizes the importance of technologies capable of removing gigatonnes of CO₂ each year, with Geological Carbon Storage (GCS) being a central component of this strategy (Panel on Climate Change) 2018; Ringrose 2020). GCS involves capturing CO₂ and injecting it into deep geological formations, where it can be safely stored over long periods. To ensure the safety of GCS, accurate monitoring of subsurface CO₂ behavior is essential, as it guarantees that CO₂ plumes remain contained and do not leak into surrounding formations (Ringrose 2023). Time-lapse seismic imaging has proven vital for tracking the migration of CO₂ plumes, yet it often fails to fully capture the complexities of multi-phase subsurface flow.

Digital Shadows (DS), which utilize machine learning-driven data assimilation techniques such as nonlinear Bayesian filtering and generative AI (Spantini, Baptista, and Marzouk 2022; Gahlot et al. 2024), offer a more detailed and reliable approach for monitoring CO₂ storage (Herrmann 2023; Gahlot et al. 2023, 2024). By incorporating uncertainty in reservoir properties like permeability, this framework improves the accuracy of CO₂ migration forecasts, including plume pressure and saturation, and thereby reduces risks for GCS projects. However, data assimilation depends on assumptions regarding reservoir properties, rock physics models linking reservoir state to seismic properties, and initial conditions. If these assumptions are inaccurate, predictions may become unreliable, which in turn can jeopardize the safety of GCS operations. One way to mitigate this risk is by augmenting the forecast ensemble used to train the neural networks responsible for the data assimilation process—mapping prior forecast samples to the posterior. In this presentation, we demonstrate that augmenting the forecast ensemble by incorporating various rock physics models mitigates the negative effects of using inaccurate models (e.g., uniform versus patchy saturation models). Additionally, we find that ensemble augmentation can enhance predictive accuracy in certain scenarios.

2 Methodology

Building on the formulation in Gahlot et al. (2024) for CO₂-plume dynamics and observations, we introduce an uncertainty-aware Digital Shadow framework, described by the following equation:

$$\begin{aligned}\mathbf{x}_k &= \mathcal{M}_k(\mathbf{x}_{k-1}, \kappa_k; t_{k-1}, t_k) \\ &= \mathcal{M}_k(\mathbf{x}_{k-1}, \kappa_k), \quad \kappa_k \sim p(\kappa) \quad \text{for } k = 1, \dots, K.\end{aligned}\tag{1}$$

Here, \mathbf{x}_k represents the time-varying spatial distribution of CO₂ saturation and pressure perturbations measured at time steps $k = 1, \dots, K$, while \mathcal{M}_k is the multi-phase fluid-flow simulation operator that evolves the state from time t_{k-1} to t_k for each time step. The symbol κ denotes the permeability distribution, which is highly heterogeneous (Ringrose 2020, 2023). Since exact permeability values are unknown, we assume a statistical distribution $p(\kappa)$ and sample from it during each timestep. The state of CO₂ saturation and reservoir pressure perturbations, \mathbf{x}_{k-1} , is then propagated forward to \mathbf{x}_k .

While reservoir simulations provide accurate modeling of CO₂ plume dynamics, relying solely on these simulations is impractical due to the inherent stochasticity of permeability and the uncertainty in rock physics models, which remain poorly constrained. To improve the accuracy of CO₂ plume forecasts, monitoring is key. Time-lapse seismic data (Lumley 2010) offer essential insights into plume behavior, and we integrate these data into our framework to enhance the characterization of the CO₂ plume. The seismic data are modeled as:

$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k; \mathcal{R}_k) + \epsilon_k, \quad \epsilon_k \sim p(\epsilon), \quad \mathcal{R}_k \sim p(\mathcal{R}) \quad \text{for } k = 1, \dots, K\tag{2}$$

where \mathcal{H}_k is the observation operator, \mathbf{y}_k is the seismic data observed at timestep k , and ϵ_k represents colored Gaussian noise added to the seismic shot records before reverse-time migration. The symbol \mathcal{R}_k denotes the rock physics model, randomly drawn from the family of Brie Saturation models (Avseth, Mukerji, and Mavko 2010), with the exponent e uniformly sampled from the range $e \sim \mathcal{U}(1, 10)$. To combine fluid-flow simulations with time-lapse seismic data, we use a nonlinear Bayesian filtering approach, incorporating neural posterior density estimation to approximate the posterior distribution of the CO₂ plume state conditioned on seismic observations. This approach merges sequential Bayesian inference with amortized neural posterior density estimation (Papamakarios et al. 2021). To estimate the posterior distribution of the CO₂ plume from seismic data, we employ Conditional Normalizing Flows (CNFs) (Gahlot et al. 2024). In this simulation-driven framework, we generate simulation pairs consisting of CO₂ plume dynamics and corresponding seismic observations (cf. equations 1 and 2) and train the CNFs on these pairs. To account for different rock physics models, we create multiple seismic images for each CO₂ plume sample, resulting in an augmented dataset.

3 Synthetic Case Study

To evaluate our approach, we use a synthetic 2D Earth model derived from the Compass model (E. Jones et al. 2012), representative of the North Sea region. We focus on a subset of this model, containing subsurface features suitable for CO₂ injection, which is discretized into a grid of 512×256 with a grid spacing of 6.25 m. To initialize the probabilistic digital shadow, we require an ensemble of potential CO₂ plume scenarios. Given the uncertainty in the permeability distribution of CO₂ storage sites, we first establish a probabilistic baseline for permeability, which we use to initialize the plume. The permeability distribution is obtained through probabilistic full-waveform inversion (Yin et al. 2024), assuming that a baseline seismic survey was conducted prior to CO₂ injection. This permeability distribution is then converted into permeability samples using the empirical relationship from (Gahlot et al. 2024).

3.1 Multi-Phase Flow Simulations

Flow simulations are carried out using the open-source tool `JutulDarcy` [JutulDarcy.jl](#) (Møyner, Bruer, and Yin 2023). Initially, the reservoir is filled with brine, and supercritical CO₂ is injected at a rate of 0.0500 m³/s over a period of 1920 days. The injection depth is approximately 1200 m. The simulation is run over four time-lapse intervals, t_k , producing predicted CO₂ saturation for each of the $N = 128$ ensemble members at each timestep.

3.2 Augmented Seismic Simulations

The outputs from the 128 flow simulations are converted into changes in subsurface acoustic properties using 10 different exponents of the Brie Saturation model (Avseth, Mukerji, and Mavko 2010), thus augmenting the data tenfold by generating different acoustic changes corresponding to each flow simulation. Seismic surveys are conducted with 8 receivers and 200 sources, using a dominant frequency of 15 Hz and a recording duration of 1.8 seconds. 28 dB SNR colored Gaussian noise is added to the shot records. Nonlinear wave simulations and imaging are performed using the open-source package `JUDI.jl` (Witte et al. 2019; Louboutin et al. 2023).

3.3 CNF Training

The 1280 ensemble members are used as training pairs, consisting of forecasted CO₂ plumes and corresponding seismic observations. We train a Conditional Normalizing Flow (CNF) using the open-source package `InvertibleNet-`

`works.jl` (Orozco et al. 2024). The network weights ϕ are optimized by minimizing the following objective over 120 epochs using the ADAM optimizer (Kingma and Ba 2014):

$$\hat{\phi} = \underset{\phi}{\operatorname{argmin}} \frac{1}{M} \sum_{m=1}^M \left(\frac{\|f_{\phi}(\mathbf{x}^{(m)}; \mathbf{y}^{(m)})\|_2^2}{2} - \log |\det(\mathbf{J}_{f_{\phi}}^{(m)})| \right). \quad (3)$$

where \mathbf{J} is the Jacobian of the network f_{θ} with respect to its input, and M is the number of training samples. For further details, we refer to (Gahlot et al. 2024).

4 Results

The performance of our enhanced Digital Shadow framework, incorporating rock physics, is shown in figure 1. The top row in each figure presents, from left to right, the ground truth (GT) CO₂ plume, the conditional mean of the posterior samples, and a sample from the posterior. The bottom row displays, from left to right, the seismic observation, the error between the conditional mean and the ground truth, and the uncertainty. As shown in figure 1a, the non-augmented DS performs well when the seismic observation matches the correct rock physics model, with the conditional mean closely approximating the GT. However, performance degrades when the seismic observation is based on a different, unknown rock physics model, as seen in figure 1b, where the conditional mean deviates from the GT, leading to increased error and uncertainty. Augmenting the ensemble improves generalization when seismic observations are based on an unknown rock physics model, as demonstrated in figure 1c. In this case, the conditional mean is closer to the GT, and both error and uncertainty are reduced compared to the non-augmented approach.

5 Conclusions

This study demonstrates that uncertainties in permeability and rock physics models can significantly impact the accuracy of CO₂ plume predictions. By augmenting the forecast ensemble, DS is able to account for unknown rock physics models (at least within the family of Brie saturation models where the exponent e is not specified) and fluid-flow properties. The enriched dataset improves the fidelity of CO₂ plume forecasts, thereby enhancing the reliability of GCS monitoring.

6 Acknowledgement

This research was carried out with the support of Georgia Research Alliance, partners of the ML4Seismic Center and in part by the US National Science Foundation grant OAC 2203821. The overall readability is enhanced using ChatGPT 4.

References

- Avseth, Per, Tapan Mukerji, and Gary Mavko. 2010. *Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk*. Cambridge university press.
- E. Jones, C., J. A. Edgar, J. I. Selva, and H. Crook. 2012. "Building Complex Synthetic Models to Evaluate Acquisition Geometries and Velocity Inversion Technologies." In *74th EAGE Conference and Exhibition Incorporating EUROPEC 2012*, cp-293. <https://doi.org/https://doi.org/10.3997/2214-4609.20148575>.
- Gahlot, Abhinav Prakash, Huseyin Tuna Erdinc, Rafael Orozco, Ziyi Yin, and Felix J. Herrmann. 2023. "Inference of CO₂ Flow Patterns – a Feasibility Study." <https://doi.org/10.48550/arXiv.2311.00290>.
- Gahlot, Abhinav Prakash, Rafael Orozco, Ziyi Yin, and Felix J. Herrmann. 2024. "An Uncertainty-Aware Digital Shadow for Underground Multimodal CO₂ Storage Monitoring." <https://doi.org/10.48550/arXiv.2410.01218>.
- Herrmann, Felix J. 2023. "President’s Page: Digital Twins in the Era of Generative AI." *The Leading Edge* 42 (11): 730–32.
- Kingma, Diederik P., and Jimmy Ba. 2014. "Adam: A Method for Stochastic Optimization." *CoRR* abs/1412.6980. <https://api.semanticscholar.org/CorpusID:6628106>.
- Louboutin, Mathias, Philipp Witte, Ziyi Yin, Henryk Modzelewski, Kerim, Carlos da Costa, and Peterson Nogueira. 2023. "Slimgroup/JUDI.jl: V3.2.3." Zenodo. <https://doi.org/10.5281/zenodo.7785440>.
- Lumley, David. 2010. "4D Seismic Monitoring of CO₂ Sequestration." *The Leading Edge* 29 (2): 150–55.
- Møyner, Olav, Grant Bruer, and Ziyi Yin. 2023. "Sintefmath/JutulDarcy.jl: V0.2.3." Zenodo. <https://doi.org/10.5281/zenodo.7855628>.
- Orozco, Rafael, Philipp Witte, Mathias Louboutin, Ali Siahkoohi, Gabrio Rizzuti, Bas Peters, and Felix J. Herrmann. 2024. "InvertibleNetworks.jl: A Julia Package for Scalable Normalizing Flows." *Journal of Open Source Software* 9 (99): 6554. <https://doi.org/10.21105/joss.06554>.

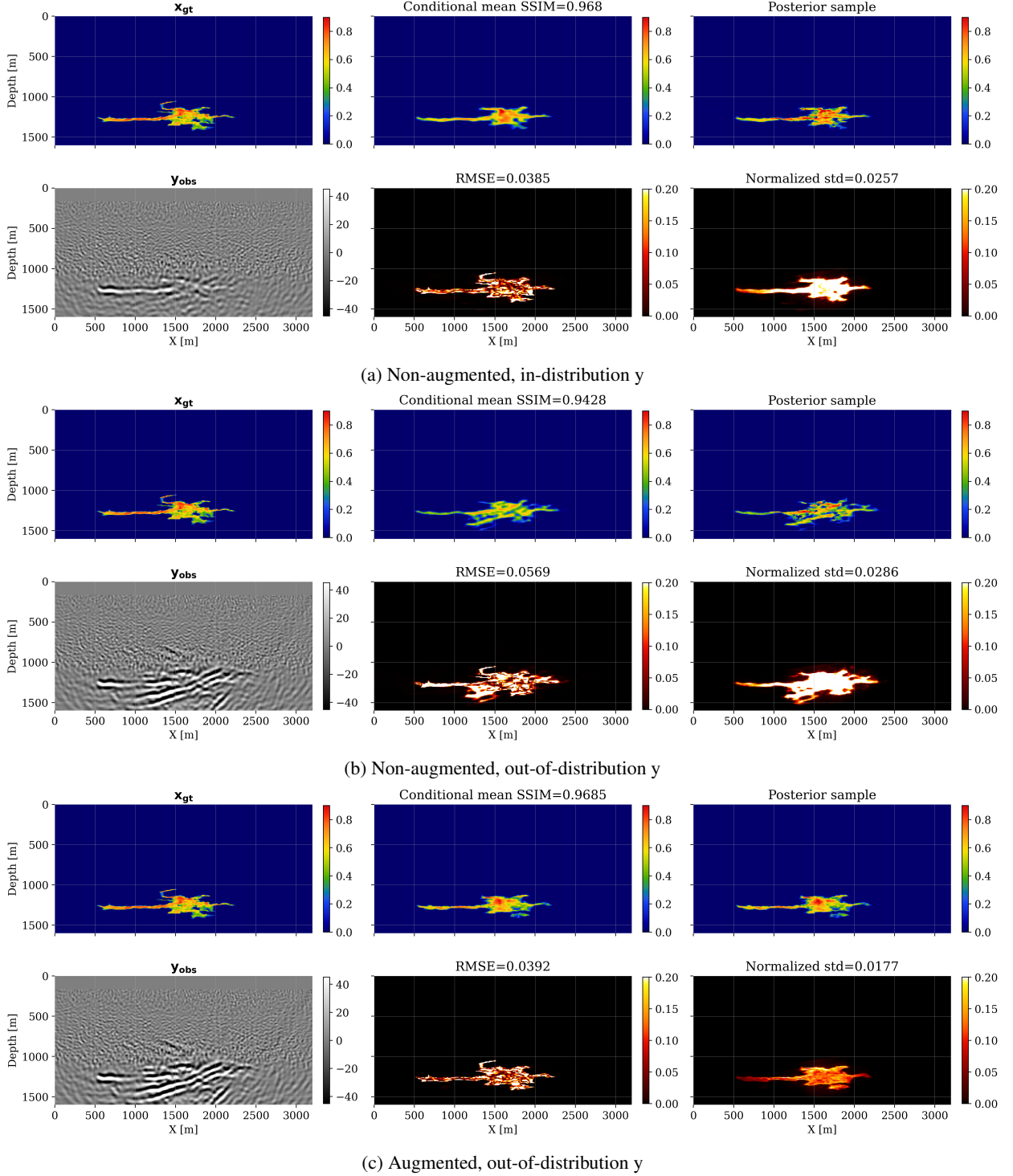


Figure 1: Inference results of digital shadow at $k = 1$ (a) without data augmentation and known rock physics model for seismic observation (b) without data augmentation and unknown rock physics model for seismic observation (c) with data augmentation and unknown rock physics model for seismic observation.

- Panel on Climate Change), IPCC (Intergovernmental. 2018. *Global Warming of 1.5°C. An IPCC Special Report on the Impacts of Global Warming of 1.5°C Above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*. ipcc Geneva.
- Papamakarios, George, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. 2021. “Normalizing Flows for Probabilistic Modeling and Inference.” *J. Mach. Learn. Res.* 22 (1).
- Ringrose, Philip. 2020. *How to Store CO₂ Underground: Insights from Early-Mover CCS Projects*. Vol. 129. Springer.
- . 2023. *Storage of Carbon Dioxide in Saline Aquifers: Building Confidence by Forecasting and Monitoring*. Society of Exploration Geophysicists.
- Spantini, Alessio, Ricardo Baptista, and Youssef Marzouk. 2022. “Coupling Techniques for Nonlinear Ensemble Filtering.” *SIAM Review* 64 (4): 921–53.
- Witte, Philipp A., Mathias Louboutin, Navjot Kukreja, Fabio Luporini, Michael Lange, Gerard J. Gorman, and Felix J. Herrmann. 2019. “A Large-Scale Framework for Symbolic Implementations of Seismic Inversion Algorithms in Julia.” *Geophysics* 84 (3): F57–71. <https://doi.org/10.1190/geo2018-0174.1>.
- Yin, Ziyi, Rafael Orozco, Mathias Louboutin, and Felix J Herrmann. 2024. “WISE: Full-Waveform Variational Inference via Subsurface Extensions.” *Geophysics* 89 (4): 1–31.