Rock Physics-Sensitivity-Aware Digital Shadow for CO$_2$ Storage Monitoring

SUMMARY

Geological Carbon Storage (GCS) requires precise CO$_2$ plume monitoring to track potential leakage pathways, but time-lapse seismic imaging alone is insufficient. A recently developed Digital Shadows (DS) enhances forecasting using machine learning and Bayesian inference, yet their accuracy depends on assumed rock physics models, the mismatch of which can lead to unreliable predictions. Augmenting DS training with multiple rock physics models mitigates errors but averages over uncertainties, obscuring their sources. To overcome this, we introduce context-aware sensitivity analysis inspired by Amortized Bayesian Inference (ABI), allowing DS to learn explicit dependencies between seismic data, CO$_2$ saturation, and rock physics models. This enables real-time adaptation to different rock physics assumptions at inference time rather than relying on costly retraining, thereby enhancing interpretability and decision-making for safer, more reliable CO$_2$ storage.

INTRODUCTION

The critical challenge of reducing global greenhouse gas emissions to mitigate anthropogenic climate change calls for the need for engineered geological carbon storage (GCS). Since the industrial revolution, the extensive combustion of fossil fuels has significantly increased atmospheric CO2 ˜ concentrations, contributing to global warming (Ringrose, 2020). Despite growing efforts to transition towards renewable energy sources, the current global energy system remains heavily reliant on fossil fuels, particularly in power generation, manufacturing, and transportation (International Energy Agency (IEA), 2016). Given the intermittency of renewable energy and the scale of emissions reduction required, an integrated approach combining multiple mitigation strategies is necessary. GCS plays a vital role in this transition by providing a means to decarbonize existing energy infrastructure and industrial processes while enabling negative emissions when coupled with bioenergy (IPCC special report, 2018; Ringrose, 2020). Furthermore, GCS facilitates a more cost-effective and timely reduction in emissions compared to relying solely on renewable energy expansion (International Energy Agency (IEA), 2016; Peters et al., 2017; Ringrose, 2020). However, to align with the targets set by the Paris Agreement ("United nations climate change conference," 2015), the deployment of GCS must be significantly scaled up, necessitating advancements in geological CO$_2$ storage technologies and the expansion of large-scale storage facilities (Global CCS Institute (GCCSI), 2019). This necessitates precise monitoring of subsurface CO$_2$ dynamics to ensure that CO$_2$ plumes remain confined within the intended storage formations, preventing migration into adjacent strata (Ringrose, 2023) and ultimately avoiding release into the atmosphere. While time-lapse seismic imaging is essential for monitoring CO$_2$ plume migration, it often lacks the resolution to fully characterize the complexities of multi-phase subsurface flow.

Digital Shadows (DS), powered by machine learning-driven data assimilation techniques such as nonlinear Bayesian filtering and generative AI (Gahlot, Orozco, Yin, & Herrmann, 2024; Spantini, Baptista, & Marzouk, 2022), offer a high-fidelity approach to characterize subsurface CO$_2$ flow (Gahlot, Erdinc, Orozco, Yin, & Herrmann, 2023; Gahlot et al., 2024; Herrmann, 2023). By incorporating uncertainty in reservoir properties such as permeability, DS provides uncertainty-aware CO$_2$ migration forecasts, including predictions of plume pressure and saturation, thereby reducing risks in geological carbon storage (GCS) projects by enabling decision-making.

However, the accuracy of these forecasts depends on assumptions regarding reservoir properties, rock physics models, and initial conditions. If these assumptions are inaccurate, predictions can become unreliable, compromising GCS safety. Gahlot & Herrmann (2025) showed that when a DS is trained assuming a particular rock physics model and evaluated with a different rock physics, it outputs an incorrect plume prediction. To address this challenge, Gahlot & Herrmann (2025) proposed augmenting the forecast ensemble used for training neural networks in the data assimilation process. By incorporating multiple rock physics models ranging from patchy to uniform saturation, their approach mitigates the impact of model misspecification and enhances predictive accuracy in key scenarios.

While this approach marginalizes over plausible rock physics models, it loses its capability for explicit reasoning. By producing a marginal posterior, the plume predictions become averaged across all possible rock physics models, resulting in broader uncertainty bounds, which can obscure the true source of uncertainty and complicate decision-making. Without an explicit treatment of rock physics variability, it becomes unclear whether uncertainties arise due to seismic noise, uncertain permeability, or incorrect assumptions about the rock physics model itself. This lack of interpretability can reduce confidence in monitoring and forecasting CO$_2$ plume behavior in geological storage settings.

To circumvent this, we need a framework that enables sensitivity analysis of the Digital Shadow over a range of rock physics models. However, training a separate network for each rock physics model is infeasible, as retraining is costly and data-intensive. Traditionally, sensitivity analyses have relied on computationally expensive refitting procedures, where a new model is trained for each specific configuration, limiting scalability and generalization in real-world monitoring scenarios. Elsemüller et al. (2024) demonstrates that Amortized Bayesian Inference (ABI) (Radev, Mertens, Voss, Ardizzone, & Köthe, 2020) provides a principled way to assess the sensitivity of

model predictions without requiring separate network training for each context. ABI not only enables rapid inference across unseen datasets but also facilitates systematic uncertainty quantification by incorporating context variables into the neural network. These context variables allow the model to learn a context-aware conditional posterior, thereby capturing the dependence of inferred parameters on different forward model assumptions. This allows for a faster and more structured approach to sensitivity analysis, where the impact of different rock physics models, in the case of DS, can be assessed directly at inference time rather than through costly retraining.

Motivated by this, we extend the approach in Gahlot & Herrmann (2025) by incorporating explicit rock physics sensitivity analysis within the DS framework. Instead of training separate models for different rock physics configurations, we introduce context variables representing the rock physics model into the training process. This allows the network to learn an explicit mapping between seismic data, CO_2 saturation, and the rock physics model, thereby improving robustness, interpretability, and generalization. By including rock physics as a context variable, we perturb seismic images corresponding to CO_2 saturation forecasts in the training ensembles in a structured manner. This ensures that the DS can dynamically adapt to different rock physics models during inference rather than producing a posterior smeared over all possible models. As a result, our approach enables not only uncertainty quantification over CO_2 plume forecasts but also a direct assessment of how different rock physics assumptions affect CO_2 plume prediction results, enhancing decision-making for geological carbon storage monitoring.

METHODOLOGY

Building upon the uncertainty-aware Digital Shadow (DS) framework introduced in (Gahlot et al., 2024), we develop a Bayesian inference-driven approach for tracking CO_2 plume evolution in geological storage reservoirs. The temporal dynamics of CO_2 saturation and pressure perturbations are modeled as:

$$\mathbf{x}_k = \mathscr{M}_k(\mathbf{x}_{k-1}, \kappa_k), \ \kappa_k \sim p(\kappa) \quad for \quad k = 1, \ldots, K. \quad (1)$$

where $\mathbf{x}_k$ represents the CO_2 saturation and pressure perturbations at time step k, governed by the multi-phase fluid flow operator $\mathscr{M}_k$. The permeability field κ, which plays a crucial role in CO_2 migration and storage efficiency, is inherently uncertain due to subsurface heterogeneity (Ringrose, 2020). To account for this uncertainty, we model permeability as a random variable drawn from a probability distribution $p(\kappa)$. This ensures that the simulation framework can generate a diverse set of plausible CO_2 migration scenarios, capturing the inherent variability of geological formations.

While fluid flow simulations provide physically consistent plume dynamics, they remain impractical without observational constraints due to the inherent stochasticity of permeability and the uncertainty in rock physics models. Time-lapse seismic imaging (Lumley, 2010) serves as a key monitoring tool, providing indirect observations of CO_2 plume migration. The corresponding observation model is defined as:

$$\mathbf{y}_k = \mathscr{H}_k(\mathbf{x}_k; \mathscr{R}_k) + \varepsilon_k, \quad \varepsilon_k \sim p(\varepsilon), \quad \mathscr{R}_k \sim p(\mathscr{R}) \quad for \quad k = 1, \ldots, K \quad (2)$$

where $\mathbf{y}_k$ represents the seismic data recorded at time t_k, $\mathscr{H}_k$ is the seismic forward operator, and ε_k is the colored Gaussian noise added to the seismic shot records before reverse-time migration and accounts for uncertainties in wave propagation and measurement errors. The term $\mathscr{R}_k$ corresponds to the rock physics transformation, which links fluid flow properties to seismic attributes. To account for uncertainty in rock physics, we sample $\mathscr{R}_k$ from a family of Brie Saturation models, with the exponent e uniformly drawn from the range $e \sim \mathscr{U}(1, 10)$. This exponent brings a context-aware data augmentation strategy that ensures that seismic images reflect a spectrum of plausible rock physics assumptions. To infer the posterior distribution of CO_2 plume states from seismic data and the context, we employ Conditional Normalizing Flows (CNFs, (Gahlot et al., 2024; Papamakarios, Nalisnick, Rezende, Mohamed, & Lakshminarayanan, 2021)) within a simulation-based inference framework utilizing the training pairs of CO_2 plume states and corresponding seismic images by varying both permeability and rock physics models (cf. equations **??** and **??**). This results in a richly augmented training set, where multiple seismic realizations are generated along with their contexts for each plume sample. The CNF approximates the posterior distribution $p(\mathbf{x}_k|\mathbf{y}_k)$, allowing for real-time Bayesian updates as new seismic data become available. This approach marries sequential Bayesian inference with neural posterior density estimation, leveraging deep generative models for rock physics context-aware CO_2 plume forecasting.

SYNTHETIC CASE STUDY

To assess the proposed methodology, we utilize a synthetic 2D Earth model derived from the Compass model (E. Jones, A. Edgar, I. Selvage, & Crook, 2012), which is representative of geological formations in the North Sea region. A subset of this model, containing key subsurface structures suitable for CO_2 injection, is selected and discretized into a computational grid of 512×256 with a spatial resolution of 6.25 m. The initialization of the DS requires an ensemble of potential CO_2 plume scenarios, which depend on the inherent uncertainty in the permeability distribution of the storage reservoir. To account for this variability, we establish a probabilistic baseline velocity model, inferred through full-waveform inversion (Yin, Orozco, Louboutin, & Herrmann, 2024) under the assumption of a baseline seismic survey conducted before CO_2 injection. The resulting samples of the velocity distribution are then converted into permeability samples using an empirical transformation elucidated in Gahlot et al. (2024).

Multi-Phase Flow Simulations

Flow simulations are conducted using the open-source tool Ju-

tulDarcy JutulDarcy.jl (Møyner, Bruer, & Yin, 2023). In the initial setup, the reservoir is filled with brine, and supercritical CO_2 is injected at a constant rate of 0.0500 m^3/s for 1920 days. The injection occurs at an approximate depth of 1200 m. The simulation is performed over four time-lapse intervals, denoted as t_k, generating predicted CO_2 saturation values for each of the $N = 128$ ensemble members at every timestep.

Context-Augmented Seismic Simulations

The outputs from the 128 flow simulations are translated into changes in subsurface acoustic properties through the application of 6 different contexts (exponent of the Brie Saturation model (Avseth, Mukerji, & Mavko, 2010)). This approach effectively augments the dataset by a factor of six, generating distinct acoustic changes for each seismic simulation. The seismic surveys are conducted using 8 receivers and 200 sources, with a dominant frequency of 15 Hz and a recording duration of 1.8 seconds. To simulate real-world conditions, 28 dB SNR colored Gaussian noise is added to the shot records. Nonlinear wave simulations and imaging are performed using the open-source package https://github.com/slimgroup/JUDI.jlJUDI.jl(Louboutin et al., 2023; Witte et al., 2019), after incorporating various rock physics models.

CNF Training

The training dataset consists of 768 ensemble members, each consisting of a forecasted CO_2 plume, its corresponding seismic observations, and the context. A Conditional Normalizing Flow (CNF) is trained using the open-source package InvertibleNetworks.jl (Orozco et al., 2024). To avoid retraining neural approximators for each context, we incorporate the context C into the network's amortization framework. This context C corresponds to an embedding of the Brie saturation exponent e. We integrate this context into the standard negative log-posterior objective of the CNF, represented as, $E_{\mathbf{x} \sim q_\phi(\mathbf{x}|\mathbf{y},C)} \left[-\log q_\phi(\mathbf{x}|\mathbf{y},C) \right]$, and is evaluated using the samples of the approximate surrogate distribution q_ϕ. To achieve the required amortization over a set of context variables C, which is obtained from the Brie exponent $e \sim \mathscr{U}(1,10)$, we minimize the context-aware (CA) loss, $E_{e \sim P(e)} \left[E_{(\mathbf{x} \sim q_\phi(\mathbf{x}|\mathbf{y},C)} \left[-\log q_\phi(\mathbf{x}|\mathbf{y},C) \right] \right]$, where the outer expectation is evaluated using the samples of e. Following Gahlot et al. (2024), all of these boils down to training a CNF where the network parameters ϕ are optimized by minimizing the following objective function over 300 epochs, utilizing the ADAM optimizer (Kingma & Ba, 2014):

$$\widehat{\phi} = \underset{\phi}{\arg\min} \frac{1}{M} \sum_{m=1}^{M} \left(\frac{\left\| f_\phi(\mathbf{x}^{(m)}; (\mathbf{y}^{(m)}, C^{(m)})) \right\|_2^2}{2} - \log \left| \det \left(\mathbf{J}_{f_\phi}^{(m)} \right) \right| \right)$$

(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).

RESULTS

The performance of our enhanced Digital Shadow framework, incorporating rock physics, is shown in figure **??**. The top row

Caption

in each figure presents, from left to right, the ground truth (GT) CO_2 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 **?@fig-naug-idy**, 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 **?@fig-naug-oody**, 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 **?@fig-aug-oody**. In this case, the conditional mean is closer to the GT, and both error and uncertainty are reduced compared to the non-augmented approach.

CONCLUSIONS

This study demonstrates that uncertainties in permeability and rock physics models can significantly impact the accuracy of CO_2 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 CO_2 plume forecasts, thereby enhancing the reliability of GCS monitoring.

ACKNOWLEDGEMENT

REFERENCES

10

Avseth, P., Mukerji, T., & Mavko, G. (2010). *Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk*. Cambridge university press.

E. Jones, C., A. Edgar, J., I. Selvage, J., & Crook, H. (2012). Building complex synthetic models to evaluate acquisition geometries and velocity inversion technologies. *In 74th EAGE Conference and Exhibition Incorporating EUROPEC 2012*, cp–293. doi:https://doi.org/10.3997/2214-4609.20148575

Elsemüller, L., Olischläger, H., Schmitt, M., Bürkner, P.-C., Koethe, U., & Radev, S. T. (2024). Sensitivity-aware amortized bayesian inference. *Transactions on Machine Learning Research*. Retrieved from https://openreview.net/forum?id=Kxtpa9rvM0

Gahlot, A. P., Erdinc, H. T., Orozco, R., Yin, Z., & Herrmann, F. J. (2023, October). Inference of CO2 flow patterns – a feasibility study. doi:https://doi.org/10.48550/arXiv.2311.0029010.48550/arXiv.2311:00290

Gahlot, A. P., & Herrmann, F. J. (2025). Enhancing robustness of digital shadow for CO2 storage monitoring with augmented rock physics modeling. Retrieved from https://arxiv.org/abs/2502.07171

Gahlot, A. P., Orozco, R., Yin, Z., & Herrmann, F. J. (2024). An uncertainty-aware digital shadow for underground multimodal CO2 storage monitoring. doi:https://doi.org/10.48550/arXiv.2410.0121810.48550/arXiv.2410.01218

Global CCS Institute (GCCSI). (2019). GCCSI CO$_2$RE database: 2019. Retrieved from https://co2re.co

Herrmann, F. J. (2023). President's page: Digital twins in the era of generative AI. *The Leading Edge*, *42*(11), 730–732.

International Energy Agency (IEA). (2016). 20 years of carbon capture and storage: Accelerating future deployment. Retrieved from https://www.iea.org/publications

IPCC special report. (2018). Global warming of 1.5 c. *An IPCC Special Report on the Impacts of Global Warming of*, *1*(5), 43–50.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *CoRR*, *abs/1412.6980*. Retrieved from https://api.semanticscholar.org/CorpusID:6628106

Louboutin, M., Witte, P., Yin, Z., Modzelewski, H., Kerim, Costa, C. da, & Nogueira, P. (2023, March). Slimgroup/JUDI.jl: v3.2.3. Zenodo. doi:https://doi.org/10.5281/zenodo.778544010.5281/zenodo.7785440

Lumley, D. (2010). 4D seismic monitoring of CO 2 sequestration. *The Leading Edge*, *29*(2), 150–155.

Møyner, O., Bruer, G., & Yin, Z. (2023, April). Sintefmath/JutulDarcy.jl: v0.2.3. Zenodo. doi:https://doi.org/10.5281/zenodo.785562810.5281/zenodo.7855628

Orozco, R., Witte, P., Louboutin, M., Siahkoohi, A., Rizzuti, G., Peters, B., & Herrmann, F. J. (2024). InvertibleNetworks.jl: A julia package for scalable normalizing flows. *Journal of Open Source Software*, *9*(99), 6554. doi:https://doi.org/10.21105/joss.0655410.21105/joss.06554

Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., & Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.*, *22*(1).

Peters, G. P., Andrew, R. M., Canadell, J. G., Fuß, S., Jackson, R. B., Korsbakken, J. I., ... Nakicenovic, N. (2017). Key indicators to track current progress and future ambition of the paris agreement. *Nature Climate Change*, *7*(2), 118–122. doi:https://doi.org/10.1038/nclimate320210.1038/nclimate3202

Radev, S. T., Mertens, U. K., Voss, A., Ardizzone, L., & Köthe, U. (2020). BayesFlow: Learning complex stochastic models with invertible neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(4), 1452–1466.

Ringrose, P. (2020). *How to store CO$_2$ underground: Insights from early-mover CCS projects* (Vol. 129). Springer.

Ringrose, P. (2023). *Storage of carbon dioxide in saline aquifers: Building confidence by forecasting and monitoring*. Society of Exploration Geophysicists.

Spantini, A., Baptista, R., & Marzouk, Y. (2022). Coupling techniques for nonlinear ensemble filtering. *SIAM Review*, *64*(4), 921–953.

United nations climate change conference. (2015). Paris, France: United Nations Framework Convention on Climate Change (UNFCCC).

Witte, P. A., Louboutin, M., Kukreja, N., Luporini, F., Lange, M., Gorman, G. J., & Herrmann, F. J. (2019). A large-scale framework for symbolic implementations of seismic inversion algorithms in julia. *Geophysics*, *84*(3), F57–F71. doi:https://doi.org/10.1190/geo2018-0174.110.1190/geo2018-0174.1

Yin, Z., Orozco, R., Louboutin, M., & Herrmann, F. J. (2024). WISE: Full-waveform variational inference via subsurface extensions. *Geophysics*, *89*(4), 1–31.