

A physics-guided neural network-based inversion framework for geophysical monitoring of geological CO₂ storage

Zeyu Wang

M.Sc. Thesis — Applied Earth Sciences

Delft University of Technology

Faculty of Civil Engineering and Geosciences

Supervisors: Eric V., Jing S.

December 2025

Delft University of Technology
Faculty of Civil Engineering and Geosciences
Department of Geoscience and Engineering

Applied Earth Science Section

<https://www.tudelft.nl/ceg>

M.Sc. Thesis

**A physics-guided neural
network-based inversion framework
for geophysical monitoring of
geological CO₂ storage**

by

Zeyu Wang

This work was performed in:

Applied Earth Sciences, Department of Geoscience,
Faculty of CEG, Delft University of Technology, The Netherlands.

DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF GEOSCIENCE

The undersigned hereby certify that they have read and recommend to the Faculty of Civil Engineering and Geosciences for acceptance a thesis entitled
“A physics-guided neural network-based inversion framework for geophysical monitoring of geological CO₂ storage”
by **Zeyu Wang** in partial fulfilment of the requirements for the degree of
Master of Science.

Dated: 03/12/2025

Chairman: Eric Verschuur

Advisor: Jing Sun

Committee Rouhi Farajzadeh

Member:

Abstract

Reliable monitoring of geologic carbon storage (GCS) requires converting time-lapse seismic data into quantitative maps of fluid movement. This thesis introduces a neural-network-based framework that embeds the seismic and rock-physics mappings during training via a physics-guided objective. With a differentiable, angle-aware geophysical forward model derived from linearized AVO, the network predicts the non-wetting-phase saturation field. The physics-guided loss computation is in line with geophysical forward operator used to synthesize the inputs—tie learning to explicit geophysical assumptions and improve stability and interpretability. The goal is fast inference of saturation changes from angle gathers while retaining interpretability and stability.

We test the framework on synthetic time-lapse data generated from known saturation fields and compare each predicted frame with ground truth. The approach recovers plume shape and migration over time, with a conservative bias at sharp fronts (slightly softer edges and reduced interior contrast), as expected from band-limited imaging and smoothness priors. Overall, the results indicate that the framework is practical for monitoring tasks that prioritize plume localization and extent. The trained model can be deployed to other sites with similar geology, acquisition, and processing; for different settings, site-specific fine-tuning or retraining is recommended. The same workflow extends to related subsurface storage problems (e.g., hydrogen) after updating the rock-physics parameters and training data.

Acknowledgment

I am deeply grateful to my supervisors, Eric and Jing, for their guidance and encouragement. The way they shaped this topic together last summer still stays with me: Eric constantly opened space for new possibilities, and Jing consistently helped turn ideas into workable plans. Their creativity, rigor, and kindness made this research a professional experience I will always cherish. Beyond academics, their care helped me grow in confidence and feel at home in the academic world.

I would also like to thank Dr. Rouhi Farajzadeh for kindly serving on my thesis committee.

My heartfelt thanks go to my family and close friends for their constant support. Living abroad alone was not always easy; during difficult nights and frustrating problems, I could always count on a late call and a steady voice. Because of you, I had the courage to explore—tasting new food and seeing beautiful places across Europe, spending three months on exchange in Ghana, and discovering a new hobby in bouldering. Our talks about life choices, philosophy, movies, and musicals enriched my perspective and kept me balanced.

I am equally grateful to everyone in the MSc Applied Earth Sciences program. I learned a great deal from many generous classmates, mentors, and staff. Your feedback, collaboration, and everyday kindness made the journey rewarding and memorable.

To all who supported me—thank you.

Contents

Abstract	3
Acknowledgment	4
1 Introduction	7
1.1 Research background	7
1.2 Literature review	8
1.3 Thesis outline	11
1.4 Research question formulation	12
2 Data and forward modeling	13
2.1 Ground-truth saturation simulation	13
2.2 Base-survey elastic properties computation	14
2.3 Differentiable forward model	15
3 Methodology	24
3.1 Pipeline overview	24
3.2 Neural network training strategy	26
3.3 Training Objective	26
4 Results	28
4.1 Qualitative comparison over time	28
4.2 Error characteristics	28
4.3 Seen versus unseen time steps	29
4.4 Additional quantitative indicators	29
4.5 Training dynamics	30
5 Discussion	34
5.1 Uncertainty qualification	34
5.2 Limitations	35
5.3 Potential improvements	36
5.4 Model performance with shallower architectural depth.	37

6 Conclusion	40
6.1 Summary	40
6.2 Main findings	40
6.3 Implications for monitoring	41
References	42
Appendix	48

Chapter 1

Introduction

1.1 Research background

Rapidly increasing greenhouse gas (GHG) emissions have very likely caused climate change, influencing the planet with a range of observable effects. Among GHG, carbon dioxide (CO₂) is acknowledged as the predominant one responsible for global warming caused by human actions. Therefore, it is imperative for policymakers to adopt effective strategies to mitigate climate change by reducing CO₂ emissions [32]. Geological CO₂ storage (GCS) technologies can play a strategic role in mitigation by injecting captured CO₂ into underground geological formations for secure, long-term storage. Especially for GHG emissions from hard-to-abate industrial processes, it emerges as a significant strategy in the global pursuit of achieving a net-zero society by 2050 [47]. In addition, it offers one of the few truly scalable approaches for reducing atmospheric CO₂ concentrations.

Monitoring represents one of the most critical components in GCS projects, as it minimizes the risk of unintended migration of the injected CO₂ from storage formations.[12], [32], [57] Monitoring of CO₂ typically relies on seismic surveys. However, effective monitoring of CO₂ storage sites should not only locate the CO₂ plume but also provide quantitative information on CO₂ saturation [24]. Seismic inversion has the potential to serve as a catalyst for quantitative monitoring. Despite the quantitative monitoring of CO₂, limitations exist.

To overcome the heavy computation involved in conventional time-lapse seismic inversion, the integration of deep learning can provide a state-of-the-art mitigation in the estimation of subsurface properties. Incorporating physics and proper architectures such as U-Net, ResNet, or Transformer into a data-driven process has the potential to enhance the capabilities of geologic

carbon storage monitoring. These deep learning advancements can potentially foster more efficient data processing, improve accuracy in subsurface property estimation, and enhance real-time monitoring applications.

Traditionally, deep learning models are trained on large amounts of labeled data, while the acquisition of large amounts of field data for geophysical exploration is nearly impossible. However, in this study, In this study, we avoid the need for labeled or field data by training a self-supervised model on synthetic data, guided by a linearized inversion strategy to rapidly generate angle gathers at multiple time instances. It enables neural networks to learn the general link between angle gathers and fluid saturation changes during monitoring phases.

The purpose of this study is to build a physics-guided deep learning framework and leverage DNN’s strong parameterization capability to solve the geophysical inversion problem.

1.2 Literature review

Relevant GCS projects and datasets for geological characterization

Prospective CO₂ storage intervals are placed at depths greater than roughly ~850 m, and are overlain by one or more low-permeability caprock units that limit upward migration. Storage formations are further expected to exhibit sufficient porosity and permeability to accept industrial injection rates and store large volumes. Candidate storage settings include depleted hydrocarbon fields, deep saline aquifers, basalt formations, and unmineable coal seams; among these, saline aquifers are frequently highlighted for their large storage potential and broad availability [36, 47], and their typically favorable petrophysical properties [21]. Consequently, some geological assumptions can adopted if a reservoir follows a deep saline aquifer setting, with regional analogs including the Utsira Sand at Sleipner and the Northern Lights project in the North Sea.

Because extensive field datasets are not always available for method development, many studies evaluate workflows on synthetic models designed to reflect realistic, deployable storage geometries. This practice is commonly justified by the fact that highly complex structural settings are often avoided in GCS site selection [26]. In this context, the 11th Society of Petroleum Engineers Comparative Solution Project (SPE 11) [38] provides a widely used benchmark geometry motivated by the Norwegian Continental Shelf (NCS) [36]. Following this work, it is feasible to generate angle gathers from 2D

base-survey elastic-property models and saturation change of CO₂. Consistent with prior practice, Araya-Polo et al. [4] trained seismic-to-model networks on synthetically generated seismic/velocity pairs; Liu and Zhong [31] used synthetic elastic models to supervise inversion; Kazei et al. [25] benchmarked deep-learning inversion on synthetic datasets; Simon et al. [45] trained on pre-stack data simulated from simplified subsurface models; and Um et al. [48] employed synthetic forward modeling to learn CO₂ saturation from seismic. Their methods of synthesizing data enable broad, controllable training distributions without relying on scarce labeled field data.

Time-lapse seismic inversion for saturation monitoring

Time-lapse (4D) seismic monitoring is widely used to infer changes in reservoir state (e.g., saturation or pressure) from differences between base and monitor surveys [22, 50]. After CO₂ injection into brine-saturated rocks, elastic properties and seismic velocities change, and repeated surveys can be inverted to map CO₂ migration and distribution [24, 34]. Field and laboratory evidence supports this approach: rock-core experiments indicate that P-wave velocity (V_p) is particularly sensitive to CO₂ saturation at relatively low saturations (below \sim 30–40%) [28], and Sleipner monitoring reports strong V_p changes in this regime that enable seismic tracking [29]. In terms of workflows, fluid-substitution and rock-physics modeling are commonly coupled with seismic forward modeling to predict time-lapse responses [10]; in practice, inversion to impedance changes is often used to obtain clearer quantitative signatures than direct differencing of seismic amplitudes [3].

Two major families of inversion dominate the time-lapse literature: full-waveform inversion (FWI) and amplitude-versus-offset (AVO) inversion. FWI can provide high resolution but requires repeated wave-equation solves and is sensitive to the starting model and noise, with cycle skipping as a well-known risk [46]. AVO inversion is computationally cheaper and directly targets elastic contrasts using pre-stack angle information [16]. AVO analysis is based on the Zoeppritz equations [59], and many practical workflows rely on linearized approximations for weak contrasts and small angles [37, 42]. For example, Aki and Richards [2] derived a standard linearization, while Shuey [44], Fatti et al. [15], Gray and Andersen [20], and others proposed alternative parameterizations (e.g., intercept-gradient forms [51], impedance-based forms [15], Lamé-parameter forms [20]) used in hydrocarbon and fluid detection [9]. These linearizations are most reliable for weak contrasts and moderate angles [7, 33, 43], and can degrade at larger (sub-critical) angles and stronger contrasts. To mitigate operator error, adjoint-state elastic AVO inversion [1] and data-driven linearized operators [37] have been proposed,

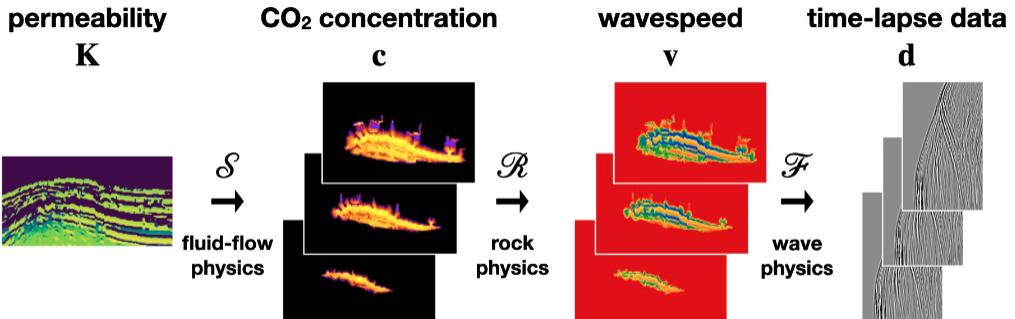


Figure 1.1: Multi-step forward modeling used to generate training pairs: saturation or permeability \rightarrow rock physics (Gassmann) \rightarrow seismic forward modeling \rightarrow synthetic seismic data [53].

typically requiring iterative updates. More broadly, several studies aim to link seismic changes to reservoir parameters (e.g., saturation or permeability) by coupling geophysics and rock physics [5], [6], [14], [19]. An example of increased integration is the end-to-end framework of Yin et al. [53], which uses pre-stack 4D seismic to invert directly for permeability, illustrating the potential of tighter coupling between seismic imaging and reservoir dynamics [52, 53]. These developments motivate approaches that retain physical interpretability while reducing the cost of repeated iterative inversion.

Neural-network-based research in seismic inversion

Recent work has explored deep learning to accelerate seismic imaging and inversion by learning mappings from seismic observations to subsurface properties. Early applications emphasize feature extraction in seismic images and the prediction of physical properties from seismic inputs [13, 56], while subsequent studies propose neural-network-based alternatives to conventional inversion workflows, often trained on synthetically generated input–output pairs [4, 25, 31, 55]. In these supervised settings, networks learn a direct mapping from seismic data to a target model (e.g., velocity or elasticity), which can improve throughput once trained but may depend strongly on the representativeness and scale of labeled pairs [27, 40].

Within this broader trend, several architectural choices recur. Convolutional neural networks (CNNs), introduced by LeCun and Bengio [30], are widely used because convolutional filters can capture localized patterns and stratigraphic textures in seismic images [58]. U-Net-style encoder–decoder networks with skip connections [41] are frequently adopted for seismic imag-

ing tasks because they preserve spatial detail while aggregating multi-scale context. For example, U-Net variants have been applied to denoising and related pre-processing tasks [11, 17], and also to forward modeling and inversion problems where inputs and outputs lie in different domains [8, 35, 54]. In the specific context of CO₂ monitoring, Um et al. [48] propose a deep-learning FWI approach in which, after training, the network estimates CO₂ saturation and associated uncertainty in near real time. Residual-network (ResNet) concepts are also used in seismic inversion to improve optimization stability and enable deeper models; attention-enhanced residual blocks have been explored to strengthen feature selectivity under noise and variability, targeting more complex geological patterns observed in seismic data.

Beyond purely spatial models, time-lapse monitoring introduces temporal structure. Sequence models such as RNNs/LSTMs are commonly used to represent temporal dependencies, but may struggle to jointly capture spatial detail and long-range temporal correlations without additional mechanisms. Transformers have therefore been explored to model long-range dependencies more effectively [18], although they typically benefit from substantial training data and may require architectural adaptations to retain spatial inductive bias. Complementing architectural advances, multiple studies emphasize physics integration to reduce non-physical solutions and improve generalization. In particular, physics-guided or physics-informed losses have been proposed to constrain learning with known forward-model structure [11], aiming to reduce reliance on spurious correlations and to stabilize training when labeled field data are scarce.

In summary, synthetic data generation is commonly used to overcome the scarcity of labeled field data in seismic ML [4, 25, 31, 45, 48]; time-lapse inversion benefits from explicitly modeling the seismic chain and rock-physics links to interpret amplitude and travel-time changes as reservoir-state changes [10, 14]; and neural networks can avoid expensive iterative inversions when supported by physical constraints and uncertainty-aware outputs [27, 48].

1.3 Thesis outline

Chapter 2 presents the data synthesis workflow, from saturation fields to seismic angle gathers via forward modeling. Chapter 3 details the network architecture and the full training strategy. Chapter 4 reports the saturation predictions. Chapter 5 analyzes uncertainty qualification, and discusses the limitations and potential improvements of the framework. Chapter 6 summarizes the main findings and conclusions.

1.4 Research question formulation

Main question. Can we build a physics-guided neural network using multi-angle, time-lapse seismic AVO information to deliver reliable saturation monitoring for GCS while preserving angle seismic information and interpretability?

Sub-questions

SQ1: Training strategy for an efficient seismic inversion. What training design enables a solver to perform seismic inversion efficiently and stably?

SQ2: Preserving angle seismic information in inversion. How can we retain and exploit the angle seismic information throughout the inversion pipeline—at input, forward operator, and loss—so that no angle seismic information averaging is performed?

SQ3: Feasibility and reliability of the prediction. How reliable is the physics-guided deep learning framework in producing operationally useful saturation maps across (i) plume localization and extent (IoU/centroid drift), (ii) boundary fidelity and amplitude consistency (NRMS), (iii) generalization to unseen time steps, and stability under noise and limited-angle conditions?

Chapter 2

Data and forward modeling

2.1 Ground-truth saturation simulation

We use the benchmark from the 11th Society of Petroleum Engineers Comparative Solution Project (11th SPE CSP)¹. The setup targets carbon storage problems with realistic geology. The public notebook runs on the Delft Advanced Research Terra Simulator (DARTS) [49].

In this study we take the project's porosity and facies description as the starting point for CO₂ migration simulation. To simulate CO₂ injection into the reservoir, we specified the area in which the saturation is only affected by CO₂ injection near the reservoir. By design, the geology is consistent with deep saline aquifers on the Norwegian Continental Shelf in the North Sea (Fig. 2.2a) [36]. Running the flow simulation produces time-lapse saturation maps, which we later use as the ground truth for forward modeling and inversion.

¹<https://spe.org/csp>; code: <https://github.com/Simulation-Benchmarks/11thSPE-CSP>

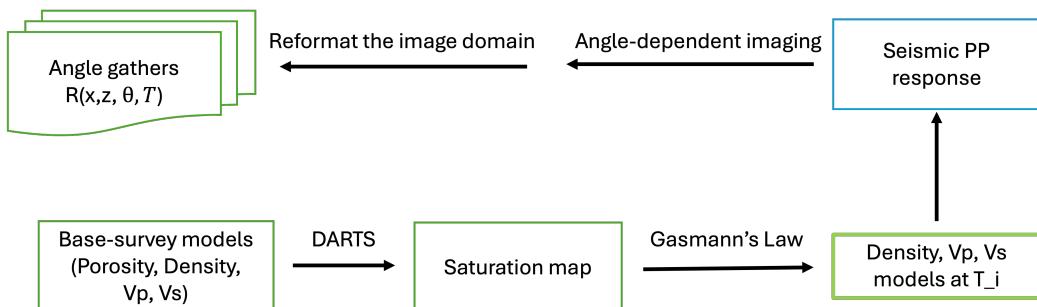


Figure 2.1: Pipeline overview of forward modeling

2.2 Base-survey elastic properties computation

The geological setting of study area in this thesis can be informed by regional analog projects and datasets. According to previous studies on the Sleipner GCS field [10], [39], the Utsira formation, the target for CO₂ injection, is a sand aquifer with an overburden depth of 700 to 900 m. Promising sites for CO₂ storage are typically deeper than \sim 850 m below ground and are overlain by one or more impermeable formations (caprock) to prevent the upward migration of CO₂. For this reason, we pad an overburden depth of 800 m atop the base-survey elastic properties (see Fig.2.3). Underneath, CO₂ is injected into a aquifer. Because velocity and density logs are not available everywhere, we build a simple and physically reasonable baseline model from porosity $\phi(x, z)$. We assume a single mineral frame (quartz-rich sand) and brine as the baseline fluid. Constants are set once (Table 2.1) and used across the model.

Table 2.1: Fixed inputs for rock and fluid used to map porosity to elastic properties.

Quantity	Symbol	Value	Unit
Mineral density	ρ_s	2500	kg m ⁻³
Mineral V_p (frame limit)	$V_{p,ma}$	3.6–3.8	km/s
Mineral V_s (frame limit)	$V_{s,ma}$	1.8–2.1	km/s
Brine density	ρ_w	\approx 1000	kg m ⁻³

We use three standard relations.

Bulk density (volumetric mixing).

$$\rho(\phi) = (1 - \phi) \rho_s + \phi \rho_f, \quad \rho_f = S_w \rho_w + (1 - S_w) \rho_{CO_2}. \quad (2.1)$$

Compressional velocity (Wyllie time average).

$$\frac{1}{V_p(\phi)} = \frac{\phi}{V_{p,fl}} + \frac{1 - \phi}{V_{p,ma}}. \quad (2.2)$$

Shear velocity (mudrock trend). With velocities in km/s,

$$V_s(\phi) = 0.862 V_p(\phi) - 1.172, \quad (2.3)$$

and we clip to a small positive lower bound at very high porosity.

These give baseline volumes (V_p, V_s, ρ) that follow regional trends. When we later test CO₂ scenarios, we keep the shear frame fixed and update only the bulk moduli using standard Gassmann fluid substitution (details in Sec. 2.3). The resulting base-survey model is shown in Fig. 2.3.

2.3 Differentiable forward model

Given baseline (Vp_1, Vs_1, ρ_1) and porosity ϕ , the network predicts non-wet saturation S_{nw} inside a reservoir. We map this to angle gathers in four steps.

(1) Fluid substitution (CO₂ effect)

We set $S_w = 1 - S_{nw}$ and perform low-frequency fluid substitution with Gassmann's relations. The workflow is:

(a) Fluid mixing. Compute effective fluid density and bulk modulus from water (w) and CO₂ (g) saturations using simple volume mixing for density and Wood's (Reuss) average for compressibility:

$$\rho_f = S_w \rho_w + S_{nw} \rho_g, \quad (2.4)$$

$$\frac{1}{K_f} = \frac{S_w}{K_w} + \frac{S_{nw}}{K_g}. \quad (2.5)$$

fluid compressibility $C_f = 1/K_f$ follows from (2.5).

(b) Recover the dry frame from the base-survey. Let base-survey (brine-saturated) properties be (V_{p0}, V_{s0}, ρ_0) with porosity ϕ and mineral (grain) moduli (K_m, μ_m) and mineral density ρ_m . First compute

$$K_{sat,0} = \rho_0 V_{p0}^2 - \frac{4}{3} \rho_0 V_{s0}^2, \quad \mu_{sat,0} = \rho_0 V_{s0}^2. \quad (2.6)$$

With base-survey fluid bulk modulus K_{f0} (e.g., brine), obtain the dry-frame bulk modulus K_{dry} via the inverse Gassmann relation:

$$K_{dry} = K_{sat,0} \frac{(\phi K_m / K_{f0}) + 1 - (K_{sat,0} / K_m)}{(\phi K_m / K_{f0}) + (K_{sat,0} / K_m) - 1}. \quad (2.7)$$

Assume the shear frame is frequency-insensitive at seismic band, so

$$\mu_{dry} = \mu_{sat,0}. \quad (2.8)$$

(c) Resaturate with the mixed fluid (Gassmann forward). Given $(K_{\text{dry}}, \mu_{\text{dry}})$, porosity ϕ , mineral modulus K_m , and mixed-fluid modulus K_f from (2.5), the saturated bulk modulus is

$$K_{\text{sat},2} = K_{\text{dry}} + \frac{\left(1 - \frac{K_{\text{dry}}}{K_m}\right)^2}{\frac{\phi}{K_f} + \frac{1-\phi}{K_m} - \frac{K_{\text{dry}}}{K_m^2}}. \quad (2.9)$$

Gassmann predicts no shear sensitivity to fluid at low frequency:

$$\mu_{\text{sat},2} = \mu_{\text{dry}}. \quad (2.10)$$

Update density by volume averaging of solid and fluid:

$$\rho_2 = (1 - \phi) \rho_m + \phi \rho_f. \quad (2.11)$$

(d) Convert to velocities.

$$V_{p,2} = \sqrt{\frac{K_{\text{sat},2} + \frac{4}{3}\mu_{\text{sat},2}}{\rho_2}}, \quad V_{s,2} = \sqrt{\frac{\mu_{\text{sat},2}}{\rho_2}}. \quad (2.12)$$

(2) Layer contrasts

Compute relative contrasts along depth:

$$\Delta V_p/V_p, \quad \Delta V_s/V_s, \quad \Delta \rho/\rho.$$

(3) Angle-dependent reflectivity (Shuey's two-term approximation)

For incidence angle α (or $s^2 = \sin^2 \alpha$), we use the common two-term linearized form

$$R(\alpha) \approx R_0 + G \sin^2 \alpha, \quad (2.13)$$

with

$$R_0 = \frac{1}{2} \left(\frac{\Delta \rho}{\rho} + \frac{\Delta V_p}{V_p} \right), \quad G = \frac{1}{2} \frac{\Delta V_p}{V_p} - 2 \left(\frac{V_s}{V_p} \right)^2 \left(\frac{\Delta \rho}{\rho} + 2 \frac{\Delta V_s}{V_s} \right). \quad (2.14)$$

These expressions give an ideal angle gather, $R_0 + G \sin^2 \alpha$, at the true reflector depth, free of wavelet and propagation effects, because migration has already placed amplitudes at the correct depths.

(4) Wavelet and gather synthesis

Before using the “real” wave propagation, we select a single vertical column at one lateral location from the 2D model and render a laterally invariant, horizontally layered Earth. We then compute multi-offset data with controlled physics via the full Zoeppritz equation on this 1.5D model (Fig. 2.4), where angle dependence comes only from contrasts in (V_p, V_s, ρ) .

The real angle-dependent image can then be compared to the ideal one. The difference is dominated by the source signature and mild residual propagation. We estimate $w(\alpha)$ by matching the ideal trend to migrated amplitudes below the overburden using a small Tikhonov-regularized least-squares fit.

Per angle, we convolve the reflectivity with a short 1D wavelet $w(\alpha)$ along depth:

$$\widehat{R}(\alpha) = r(\alpha) * w(\alpha).$$

To select proper angles for R synthesis, we perform a simple check at interfaces to confirm that Shuey’s two-term approximation is valid here. The same logs also provide the interface contrasts from which analytical amplitude trends can be predicted (Fig. 2.7). At several interfaces we compare the “ideal” coefficients from (2.14) with the measured trend after migration. Both coefficients agree for small to moderate angles ($\sin^2 \alpha \lesssim 0.4$), which is the range used for the inversion.

(5) Residuals for training input data.

We select four wavelets corresponding to four angles ($\sin^2 \alpha = 0.0, 0.1, 0.2, 0.3$). We also compute the baseline angle gathers at the same angles. To avoid filter mismatch, the same wavelet family is used to compute the baseline gather. We then form residuals (Fig. 2.8)as

$$d\widehat{R}_t(\alpha) = \widehat{R}_t(\alpha; S_{nw}) - R_0(\alpha; Vp_1, Vs_1, \rho_1).$$

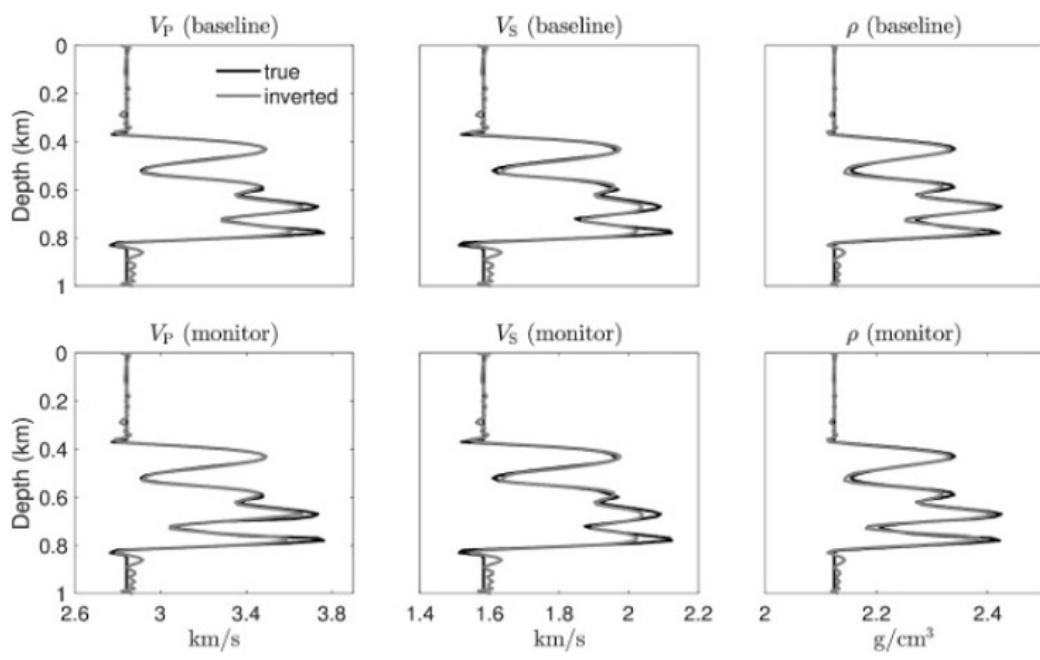
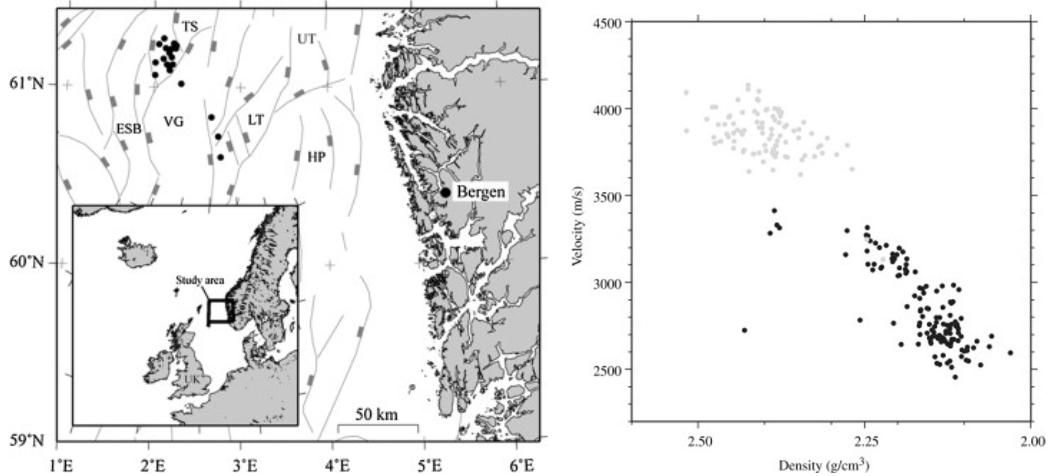


Figure 2.2: Geological context and calibration data.[36]

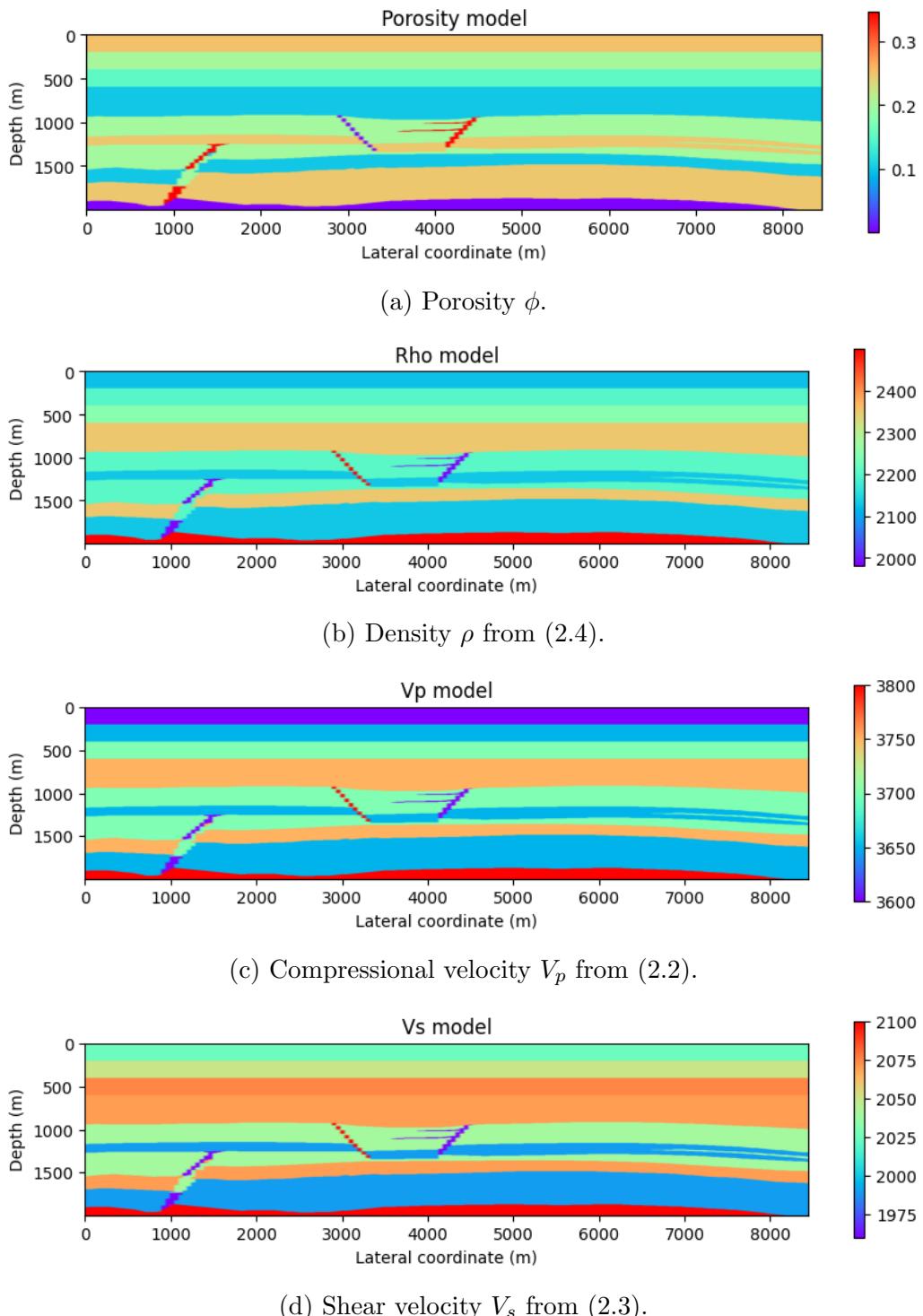


Figure 2.3: Base-survey elastic properties (4×1 layout).

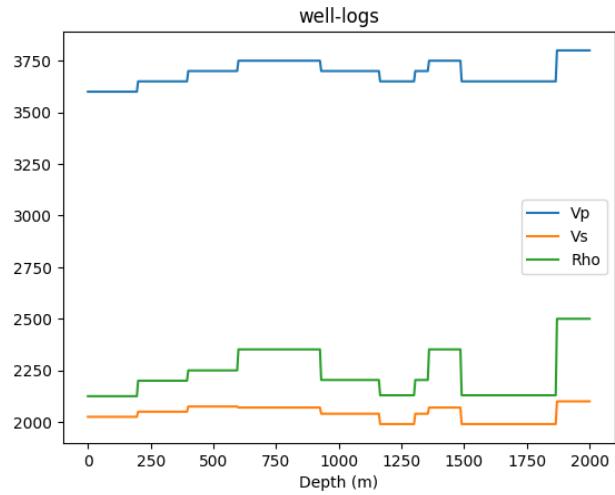


Figure 2.4: Well logs used for simple calibration and checks.

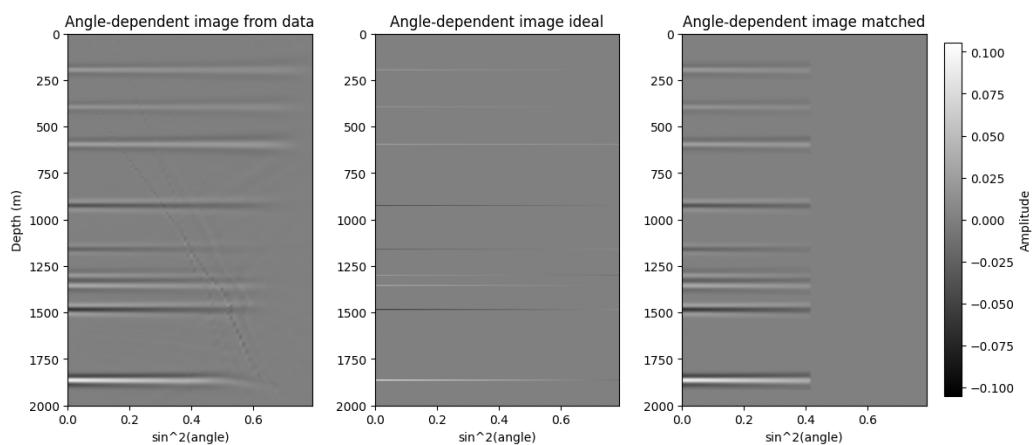


Figure 2.5: Angle-dependent images (ideal vs. real).

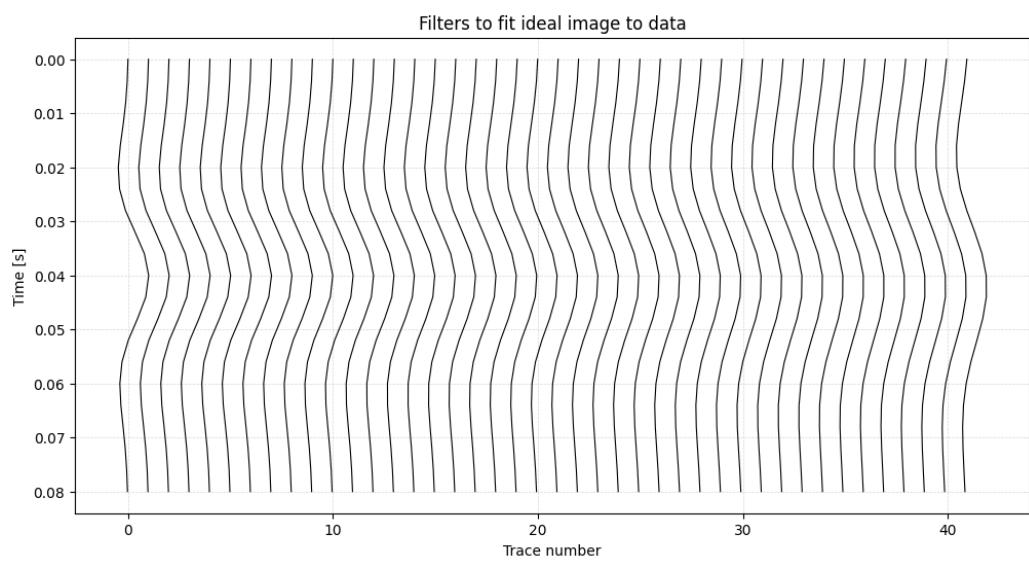
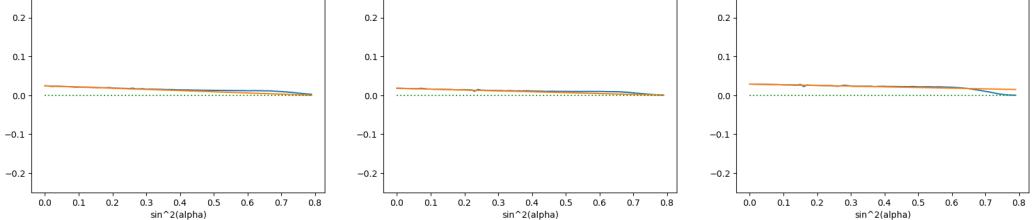
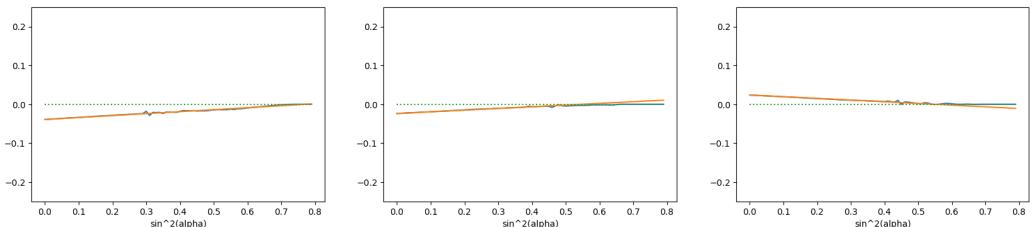


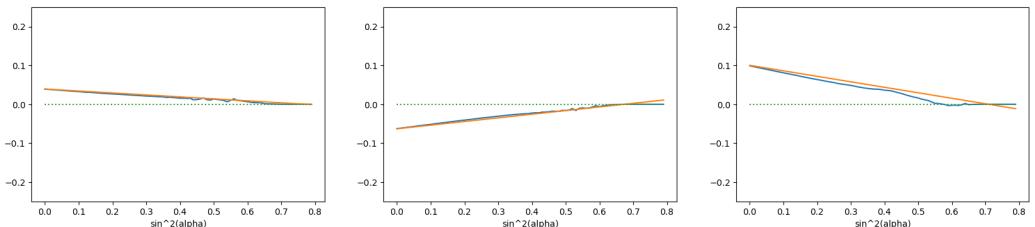
Figure 2.6: Angle-dependent shaping filter estimated by least-squares matching of ideal to migrated amplitudes.



(a) Impedance contrast 1. (b) Impedance contrast 2. (c) Impedance contrast 3.

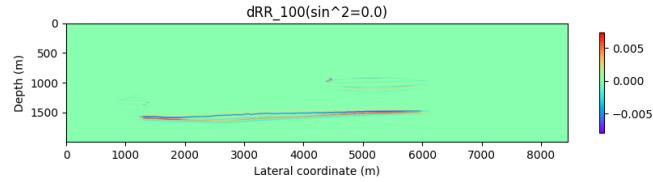


(d) Impedance contrast 4. (e) Impedance contrast 5. (f) Impedance contrast 6.

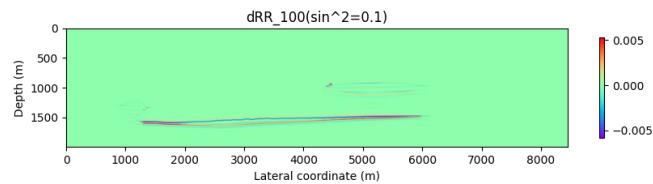


(g) Impedance contrast 7. (h) Impedance contrast 8. (i) Impedance contrast 9.

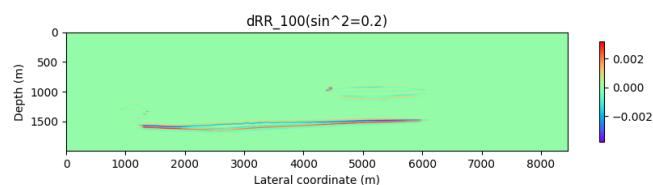
Figure 2.7: Ideal (orange) vs. measured (blue) two-term Shuey trends at each impedance-contrast interface. Densities are in kg m^{-3} ; velocities in m s^{-1} .
 (a) Found contrast at $\rho = 2125, 2200$; $V_p = 3600, 3650$; $V_s = 2025, 2050$.
 (b) Found contrast at $\rho = 2200, 2250$; $V_p = 3650, 3700$; $V_s = 2050, 2075$.
 (c) Found contrast at $\rho = 2250, 2351$; $V_p = 3700, 3750$; $V_s = 2075, 2070$.
 (d) Found contrast at $\rho = 2351, 2203$; $V_p = 3750, 3700$; $V_s = 2070, 2040$.
 (e) Found contrast at $\rho = 2203, 2129$; $V_p = 3700, 3650$; $V_s = 2040, 1990$.
 (f) Found contrast at $\rho = 2129, 2203$; $V_p = 3650, 3700$; $V_s = 1990, 2040$.
 (g) Found contrast at $\rho = 2203, 2351$; $V_p = 3700, 3750$; $V_s = 2040, 2070$.
 (h) Found contrast at $\rho = 2351, 2129$; $V_p = 3750, 3650$; $V_s = 2070, 1990$.
 (i) Found contrast at $\rho = 2129, 2500$; $V_p = 3650, 3800$; $V_s = 1990, 2100$.



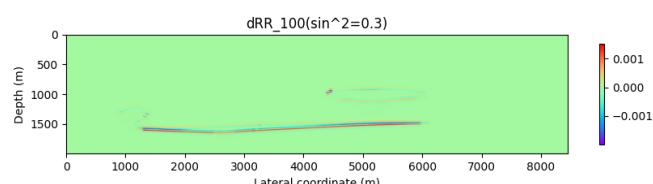
(a) $\sin^2 \alpha = 0.0$



(b) $\sin^2 \alpha = 0.1$



(c) $\sin^2 \alpha = 0.2$



(d) $\sin^2 \alpha = 0.3$

Figure 2.8: Residual angle gathers at four angles for year 100.

Chapter 3

Methodology

3.1 Pipeline overview

We develop a cost-efficient workflow to improve GCS monitoring using a neural, physics-guided inversion. The goal is to turn time-lapse seismic angle gathers into maps of saturation change and, for decision support, an updated view of permeability. The network ingests observed angle gathers d_{obs} and delivers a prediction that is directly useful for monitoring while remaining tied to simple, transparent physics.

We seek the non-wetting-phase saturation field $S_{nw}(z, x)$ from four incident angles of reflection differences $\{dR_{\theta_i}(z, x)\}_{i=1}^4$. Rather than learn a black-box mapping $dR \rightarrow S_{nw}$, the network outputs only S_{nw} . All angle-dependent reflectivity is then generated by a differentiable *physics forward model*. This cleanly separates the subsurface state from acquisition and processing effects (source wavelet, amplitude scaling, timing), and keeps the inversion interpretable.

Concretely, with inputs $\mathcal{D} = \{dR_{\theta_i}\}_{i=1}^4$,

$$S_{nw}(z, x) = f_{\Theta}(\mathcal{D}),$$

and the forward model \mathcal{P} (fluid substitution + linearized AVO + angle-dependent wavelet) maps S_{nw} and the baseline properties to predicted gathers $\widehat{dR}_{\theta_i} = \mathcal{P}(S_{nw})$. Subsequent sections define the loss and regularization used to train f_{Θ} while keeping the permeability update as a downstream, guidance product.

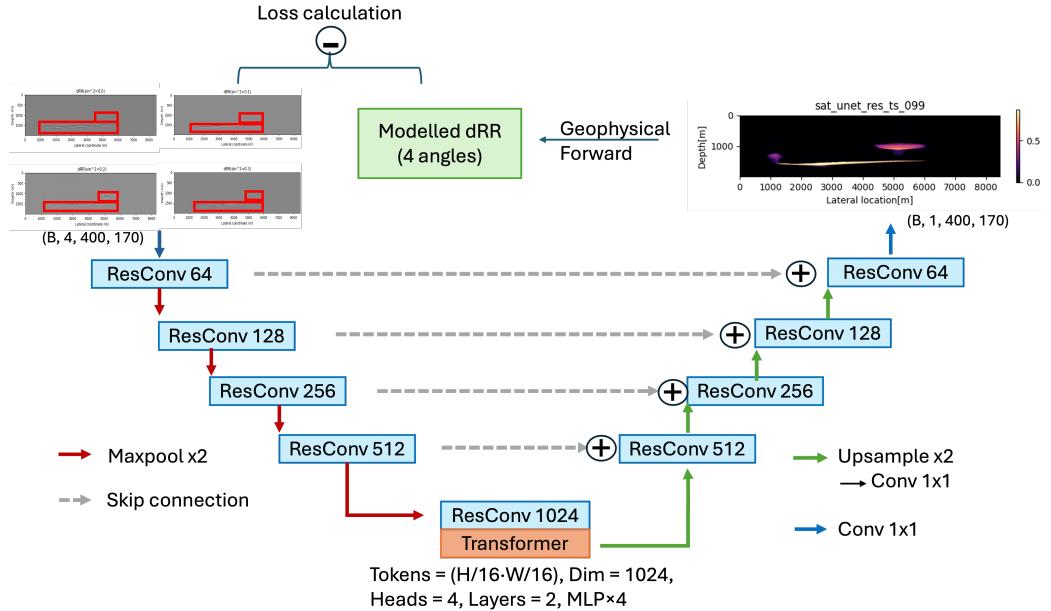


Figure 3.1: Physics-Guided neural network for Saturatation Inversion from Multi-Angle Reflectivity: UNet with Residual Blocks and Transformer Bottleneck Trained by a Physics-Driven dR Loss: the encoder uses residual conv blocks (64–512 ch) with MaxPool $\times 2$; the bottleneck (1024 ch) is processed by a 2-layer, 4-head transformer when lateral long-range structure remains under-modeled, add a light attention block at the UNet bottleneck while keeping the same physics and losses; the decoder upsamples ($\times 2 \rightarrow 1 \times 1$), concatenates mirrored skips, and outputs Snw via a 1×1 head. Training is physics-guided: predicted Snw to \hat{R}_i , with baseline removal $d\hat{R} = \hat{R}_i - \hat{R}_0$. The loss compares $d\hat{R}$ to observed dR (shared angle bins). Angles are channels $c \in \{\sin^2 \theta = 0.0, 0.1, 0.2, 0.3\}$.

3.2 Neural network training strategy

As shown in Fig.3.1, our neural network-based inversion solver is constructed on the UNet encoder-decoder architecture with flexibly switching among convolutional, Transformer, and ResNet blocks, together with physics driven. Each block consists of a series of convolution, batch normalization, and activation layers which has been used and tested in [23]. It is a model that combines the benefits of these architectures. The neural network will be trained on pairs of synthetic AVO information and CO₂ saturation change profiles. Consequently, our training establishes a mapping relationship between seismic data and CO₂ saturation changes, and once fully trained, directly estimates CO₂ saturation as a function of subsurface location.

3.3 Training Objective

We predict water saturation $S \in [0, 1]^{H \times W}$ from four relative-reflectivity channels $\mathbf{Y} \in \mathbb{R}^{4 \times H \times W}$. A differentiable forward model \mathcal{F} maps S to angle-resolved reflectivities $\widehat{\mathbf{R}} = \mathcal{F}(S)$. Let \mathbf{R}_0 denote the baseline reflectivity and define the physics residual $\mathbf{P} = \widehat{\mathbf{R}} - \mathbf{R}_0$. We compare \mathbf{P} to the measured difference $\mathbf{T} = \mathbf{Y}$ only inside porous rock by means of a porosity mask $M \in [0, 1]^{H \times W}$ (derived from a porosity volume by thresholding/dilation). We write $\mathcal{M} = \{(z, x) \mid M(z, x) > 0\}$.

Masked inner product. For tensors $A, B \in \mathbb{R}^{H \times W}$ we define the porosity-masked inner product

$$\langle A, B \rangle_M = \sum_{(z, x) \in \mathcal{M}} M(z, x) A(z, x) B(z, x). \quad (3.1)$$

Inner masked fit (per angle; stop-gradient). To handle unknown per-angle polarity and scale we perform a closed-form inner fit, choosing the sign by masked correlation and solving a masked least-squares amplitude. For channel $c \in \{1, \dots, 4\}$,

$$s_c = \text{sign}(\langle P_c, T_c \rangle_M), \quad (3.2)$$

$$a_c = \frac{\langle s_c P_c, T_c \rangle_M}{\langle P_c, P_c \rangle_M + \varepsilon}, \quad \varepsilon > 0, \quad (3.3)$$

and we form the fitted predictor

$$\widehat{\mathbf{dR}} = a \odot s \odot \mathbf{P}, \quad a = (a_c)_{c=1}^4 \text{ (treated as constant / stop-gradient)}, \quad s = (s_c)_{c=1}^4. \quad (3.4)$$

Here \odot denotes elementwise multiplication, broadcasting along spatial coordinates.

Physics data term. We penalize the porosity–masked mean–squared discrepancy between the inner–fit predictor and the target:

$$\mathcal{L}_{\text{phys}} = \frac{1}{C|\mathcal{M}|} \sum_{c=1}^C \sum_{(z,x) \in \mathcal{M}} \left(d\hat{\mathbf{R}}_c(z, x) - T_c(z, x) \right)^2, \quad C = 4. \quad (3.5)$$

Regularization: anisotropic total variation on S . We encourage spatial coherence of S with a light anisotropic TV prior ($\alpha_z \geq \alpha_x$ is typical to discourage horizontal banding):

$$\mathcal{L}_{\text{TV}} = \alpha_x \|\nabla_x S\|_1 + \alpha_z \|\nabla_z S\|_1. \quad (3.6)$$

Stripe–projection penalty. Let the angle gather data be $D \in \mathbb{R}^{Z \times \Theta}$ (depth Z by angle Θ). A depth-constant “stripe” at each angle can be written as

$$S = \mathbf{1}_Z b^\top \in \mathbb{R}^{Z \times \Theta},$$

where $\mathbf{1}_Z$ is a Z -vector of ones and $b \in \mathbb{R}^\Theta$ collects one coefficient *per angle*. We estimate b so that S captures the row-constant nuisance:

$$b = \arg \min_b \|D - \mathbf{1}_Z b^\top\|_F^2,$$

To prevent the optimizer from explaining away true signal via this nuisance component, we add a small Tikhonov penalty on the stripe coefficients,

$$\mathcal{L}_{\text{bg}} = \|b\|_2^2, \quad (3.7)$$

which biases the fitted stripe towards zero unless strongly supported by the data.

Total objective (three terms). Combining the terms, the training loss is

$$\mathcal{L}_{\text{total}} = \underbrace{\mathcal{L}_{\text{phys}}}_{\text{physics consistency}} + \lambda_{\text{TV}} \underbrace{\mathcal{L}_{\text{TV}}}_{\text{spatial smoothness}} + \lambda_{\text{bg}} \underbrace{\mathcal{L}_{\text{bg}}}_{\text{background / gain control}}, \quad (3.8)$$

where λ_{TV} and λ_{bg} balance the TV and optional stripe penalties, respectively. In practice, the amplitudes a_c in (3.3) are computed in closed form and treated as constants (detached) when backpropagating through \mathcal{L} , ensuring stable gradients with respect to S through the forward model \mathcal{F} .

Chapter 4

Results

4.1 Qualitative comparison over time

As described in Chapter 2, the input angle gathers were synthesized from ground-truth saturation fields $\{S_c(\mathbf{x}, t)\}_{t=1}^{800}$ using the forward modeling and imaging pipeline. This enables a direct, frame-wise comparison between the predicted saturation $\{\hat{S}_c(\mathbf{x}, t)\}$ and the truth. For each time step $t \in [1, 800]$ we render a side-by-side view (prediction left; ground truth right) and compile the sequence into animations to visualize temporal evolution (Fig. 4.1–4.5).

Across the sequence, the network consistently recovers the geometry of the two migrating plumes. Their nucleation zones, down-dip drift, and lateral spreading are placed at the correct locations, and the inter-plume gap is preserved. Agreement is most evident during early to mid time steps, when plumes are compact and well separated. As the plumes widen, minor softening of edges becomes visible, but the overall migration paths remain correct.

4.2 Error characteristics

Two systematic differences are observed. First, interiors of predicted plumes show reduced contrast: gradients $\|\nabla \hat{S}_c\|$ are attenuated relative to $\|\nabla S_c\|$, producing flatter plateaus. Second, plume boundaries are more diffuse, with broader transition zones than in the ground truth. Together, these effects bias the solution toward conservative, lower-variance estimates near sharp fronts. The degree of blurring increases gradually with time as plumes broaden, which is consistent with the band- and aperture-limited imaging operator and the smoothness-promoting regularization used during training.

4.3 Seen versus unseen time steps

Of the 800 time steps, 720 were used for training (seen) and 80 were held out for evaluation (unseen). On the seen set, predictions retain slightly crisper edges and stronger interior contrast, reflecting repeated exposure to similar states. On the unseen set, the two-plume geometry is still localized correctly, but edges are less sharp and the dynamic range is compressed. Thus, the generalization gap lies primarily in reduced boundary clarity and variance, not as structural misplacement. For monitoring, where plume localization is prioritized, this behavior is acceptable.

4.4 Additional quantitative indicators

To complement the visual assessment, we summarize three diagnostic indicators that are inexpensive to compute and informative for monitoring:

Plume extent (IoU). Let $\mathcal{P}_\tau(t)$ and $\mathcal{G}_\tau(t)$ be the predicted and ground-truth supports thresholded at τ (e.g., $\tau = 0.1$ in saturation units). The intersection-over-union IoU $_\tau(t) = \frac{|\mathcal{P}_\tau \cap \mathcal{G}_\tau|}{|\mathcal{P}_\tau \cup \mathcal{G}_\tau|}$ tracks extent accuracy. In our experiments, IoU is stable across time, with a small drop on unseen frames that mirrors the softening of edges.

Boundary sharpness (F1). Extracting plume boundaries (e.g., Canny edges on a smoothed map), we compute boundary precision, recall, and F1 between prediction and truth. F1 decreases mildly with time as edges diffuse, but boundary locations remain spatially aligned, consistent with correct plume positioning.

Centroid drift. For each plume, we compute the centroid in prediction and truth and report the drift norm $\|\mathbf{c}_{\text{pred}}(t) - \mathbf{c}_{\text{true}}(t)\|_2$. Drift remains small relative to plume scale across both seen and unseen sets, supporting the qualitative observation that geometry and trajectories are preserved.

These indicators provide a compact operational view: extent and trajectory remain reliable; sharpness is the principal dimension where predictions are conservative.



(a) Predicted saturation (1–25 yr). (b) Ground-truth saturation (1–25 yr).

Figure 4.1: Side-by-side evolution of CO_2 saturation change ΔS for years 1–25. (Open in Adobe Acrobat Reader to play.)

4.5 Training dynamics

The training loss decreases rapidly over the first $\sim 10\text{--}12$ epochs and then flattens into a slow, steady descent (Fig. 4.6). From the initial value (≈ 0.279) to the final value at epoch 60 (≈ 0.225), the loss is reduced by about 19%. Roughly 85–90% of that improvement is achieved by epoch 12, indicating that the model reaches a practically useful operating point early, with subsequent epochs delivering incremental refinements. Small, short-lived bumps around epochs $\sim 35\text{--}40$ suggest the optimizer explores shallow basins rather than diverging, consistent with a stable learning rate. In operational terms, this profile supports (i) early-stopping around the “knee” (epoch 10–15) when compute is limited, and (ii) a cosine or step decay schedule after epoch 12 to exploit the long tail efficiently. For completeness, we report the loss curve which would quantify generalization and guard against late-epoch overfitting; in our experiments the qualitative results (Sec. 4) are consistent with a well-behaved training trajectory.

(a) Predicted saturation (26–240 yr). (b) Ground-truth saturation (26–240 yr).

Figure 4.2: Side-by-side evolution of CO₂ saturation change ΔS for years 26–240. (Open in Adobe Acrobat Reader to play.)

(a) Predicted saturation (241–480 yr). (b) Ground-truth saturation (241–480 yr).

Figure 4.3: Side-by-side evolution of CO₂ saturation change ΔS for years 241–480. (Open in Adobe Acrobat Reader to play.)

(a) Predicted saturation (481–720 yr). (b) Ground-truth saturation (481–720 yr).

Figure 4.4: Side-by-side evolution of CO₂ saturation change ΔS for years 481–720. (Open in Adobe Acrobat Reader to play.)

(a) Predicted saturation (721–800 yr). (b) Ground-truth saturation (721–800 yr).

Figure 4.5: Side-by-side evolution of CO₂ saturation change ΔS for years 721–800. (Open in Adobe Acrobat Reader to play.)

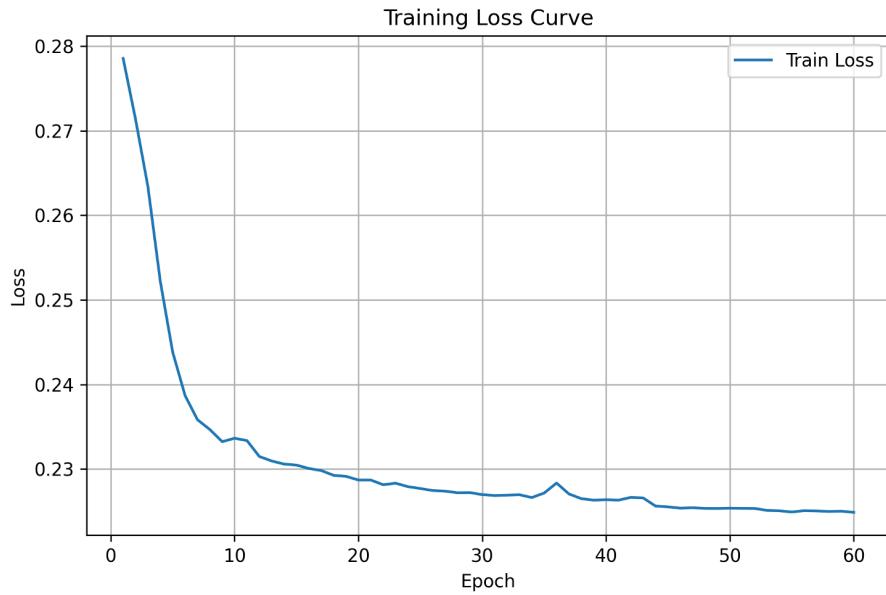


Figure 4.6: Training loss vs. epoch. Fast initial descent (epochs 1–12) followed by a long, low-slope tail. The curve suggests an early stopping point near the knee and motivates a learning-rate decay after ~ 12 epochs.

Chapter 5

Discussion

5.1 Uncertainty qualification

The two dominant artifacts noted in Chapter 4—attenuated saturation gradients inside the plumes and broadened plume boundaries—are consistent with both the physics of the imaging operator and the training regularization. The angle-gather imaging is band-limited and aperture-limited; in the $\tau-p$ continuation we mute evanescent components and propagate with a smooth background, which suppresses high wavenumbers and naturally softens sharp fronts. On the learning side, the loss includes smoothness-promoting terms (e.g., TV) and a physics term evaluated on angle images; together, these bias the model toward conservative (low-variance) estimates where the data carry less information, precisely at the thin, high-contrast rims of the plumes. None of these mechanisms displace the plumes; rather, they explain why edges appear soft and interiors slightly flat while geometry and migration paths remain correct.

We partitioned the sequence into 720 seen years for training and 80 unseen years for evaluation. Because angle gathers were synthesized from ground-truth saturation, we can quantify error directly. We report the normalized RMS difference (NRMS) between prediction $\widehat{S}_c(\cdot, \cdot, t)$ and truth $S_c(\cdot, \cdot, t)$ at each year t ,

$$\text{NRMS}(t) = 100 \times \frac{2 \|\widehat{S}_c(\cdot, \cdot, t) - S_c(\cdot, \cdot, t)\|_2}{\|\widehat{S}_c(\cdot, \cdot, t)\|_2 + \|S_c(\cdot, \cdot, t)\|_2}, \quad (5.1)$$

which is symmetric in the two fields and insensitive to uniform rescaling. The NRMS history (Fig. 5.1) shows three regimes. First, an early transient with elevated values during initial plume growth (injection continues until roughly year 50), when geometry evolves quickly and is hardest to match.

Second, a broad plateau over mid-life, where NRMS stabilizes with modest fluctuations; occasional spikes (e.g., near \sim 120, 360, and 540 yr) align with abrupt geometric updates—thin spill points, layer contacts, or coalescence/splitting events—where imaging bandwidth and learned regularity jointly limit sharpness. Third, a clear generalization gap over the *unseen* interval (years 721–800), visible as a step-up in NRMS: edges are softer and interior contrasts slightly compressed, while the two-plume geometry remains localized correctly (see $t=780$ yr in Fig. 5.4). For monitoring tasks that prioritize plume extent and trajectory, this behavior is acceptable; thickness and front sharpness should be interpreted with the stated bias in mind.

5.2 Limitations

A first limitation is the dimensionality of the training data: we synthesize 1.5D examples rather than full 2D/3D seismic images, effectively stacking 1D reflectivity responses along the lateral axis to form a pseudo-2D section; while this enables efficient training, it can under-represent genuine 2D/3D wave-propagation effects (e.g., lateral heterogeneity, diffractions, and finite-offset moveout complexity) that may influence angle gathers in practice. A second limitation concerns the AVO forward model: the current two-term linearized formulation is most reliable for small-to-moderate angles and modest impedance contrasts, and performance may degrade at larger (yet sub-critical) angles or in high-contrast settings.

A further limitation relates to the use of a lightweight porosity mask as a spatial prior. While a weak mask typically improves geometric plausibility by steering the network toward reservoir architecture (e.g., more consistent plume centroids and migration paths) and by reducing the search space in poorly illuminated zones, it can become brittle when the prior is violated. Unmodelled leakage pathways or unexpected high-permeability corridors (e.g., faults) may be under-predicted if the mask suppresses them, which can manifest as muted interiors and over-smoothed edges in those regions. To reduce this risk without discarding the prior, the mask should be applied softly, and its sensitivity should be quantified through stress tests (e.g., perturbing the mask by opening a narrow corridor and reporting the resulting change in NRMS and boundary F1). Overall, the mask is best treated as a calibration knob rather than a fixed constraint, especially at sites where leakage pathways are uncertain. Finally, global per-angle amplitude/phase corrections may not fully capture acquisition and processing imprint, and the inversion can exhibit softened plume boundaries and a stochastic spread of solutions under noise, which may limit interpretability in decision-critical

settings without additional constraints.

5.3 Potential improvements

Several practical extensions can directly address the limitations above while preserving computational efficiency. First, the method is most stable when the input angles are restricted to a moderate range (e.g., $\sin^2 \alpha \leq 0.4$), consistent with the validity regime of the two-term linearization; if impedance contrasts are stronger, robustness can be improved by extending the forward model via either the third Shuey term or full Zoeppritz reflectivity, at a modest additional cost. Second, wavelet and acquisition imprint can be handled more flexibly: while per-angle shaping already reduces systematic bias, residual mismatches can be further reduced by introducing slowly varying gain fields $g(\mathbf{x}, \theta)$ with spatial regularization and by applying a limited phase-only correction (restricted to a few samples) to absorb small timing drifts. Third, from an operational standpoint, even when plume edges are somewhat softened, accurate plume localization is often sufficient to track plume extent (a key containment metric), and this can be strengthened by adding an explicit reservoir-scale mass-balance check that is simple to audit. Specifically, we define the bulk CO₂ volume in place over the domain Ω at time t as

$$V_{\text{CO}_2}(t) = \left(\int_{\Omega} \phi(\mathbf{x}) S_{nw}(\mathbf{x}, t) d\mathbf{x} \right) \times B,$$

where ϕ is porosity and B is an optional formation-volume factor for unit consistency, and we encourage consistency with cumulative injection using

$$\mathcal{L}_{\text{mass}}(t) = \lambda_{\text{mb}} [V_{\text{CO}_2}(t) - V_{\text{inj}}^{\text{cum}}(t) + V_{\text{prod}}^{\text{cum}}(t)]_+^2, \quad [x]_+ = \max(x, 0).$$

This penalizes solutions where the inferred in-place CO₂ exceeds the net injected volume (injected minus any produced/removed volumes) and can be supplemented by weak temporal priors: during injection $V_{\text{CO}_2}(t+1) \geq V_{\text{CO}_2}(t)$ (allowing minor deviations for noise), while after shut-in (assuming no leakage) $V_{\text{CO}_2}(t+1) \approx V_{\text{CO}_2}(t)$ within imaging error; if leakage is suspected, $V_{\text{CO}_2}(t)$ should decrease and always satisfy $V_{\text{CO}_2}(t) \leq V_{\text{inj}}^{\text{cum}}(t) - V_{\text{prod}}^{\text{cum}}(t)$. Overall, these reservoir-scale balance terms complement the physics-guided loss by anchoring the inversion to an auditable deterministic constraint and, through the adopted rock-physics links, helping convert inferred velocity changes into stable saturation updates.

In terms of deployment, the pre-trained model can be used directly when the target application closely matches this study (similar geology, acquisition,

and processing). For other sites, a site-specific model should first be trained (or adapted) using representative synthetic data and rock physics; the resulting weights can then serve as an efficient initialization for subsequent monitoring inversions at that site. The same workflow is also transferable to subsurface hydrogen storage, provided the training data generation and rock-physics links are adjusted to hydrogen–brine systems prior to deployment.

In addition, the current evaluation would benefit by comparing with predictions generalized by near-zero angle gather trained model. This will explicitly quantify the incremental value of multi-angle seismic information and to isolate failure modes that persist even without angular information.

5.4 Model performance with shallower architectural depth.

Reducing architectural depth (fewer encoder/decoder stages or residual blocks) led to systematic performance losses relative to the base survey. Quantitatively, we observed higher validation NRMS and angle-consistency residuals, and larger plume extent errors. Qualitatively, saturation maps exhibited softened boundaries, attenuated thin features, and reduced interior contrast—consistent with a smaller effective receptive field and lower representational capacity. Training dynamics also indicated underfitting: the training loss plateaued earlier, and the gap between training and validation curves narrowed without corresponding accuracy gains. Overall, the results suggest that the depth used in this thesis is near the minimum required to preserve plume geometry and temporal coherence under the adopted physics-guided loss.

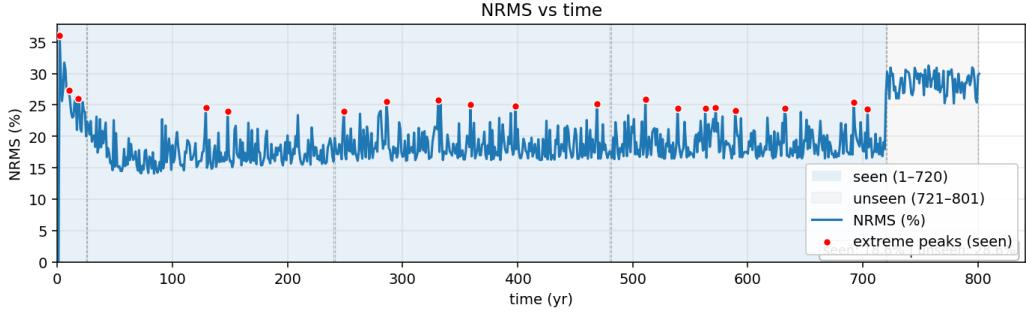


Figure 5.1: NRMS history over 800 years. The shaded region indicates the training (seen) window; the rightmost segment is the unseen evaluation window. Early years show an elevated transient tied to rapid plume growth; the mid sequence stabilizes with intermittent spikes associated with abrupt geometric updates; the unseen interval exhibits a step-up in NRMS while preserving plume localization.

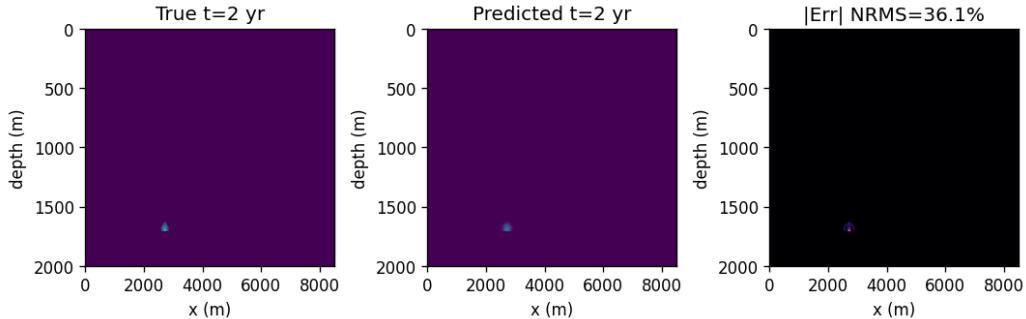


Figure 5.2: Representative frame at $t = 2$ yr. **Left:** ground truth ΔS . **Middle:** prediction. **Right:** absolute error with NRMS indicated.

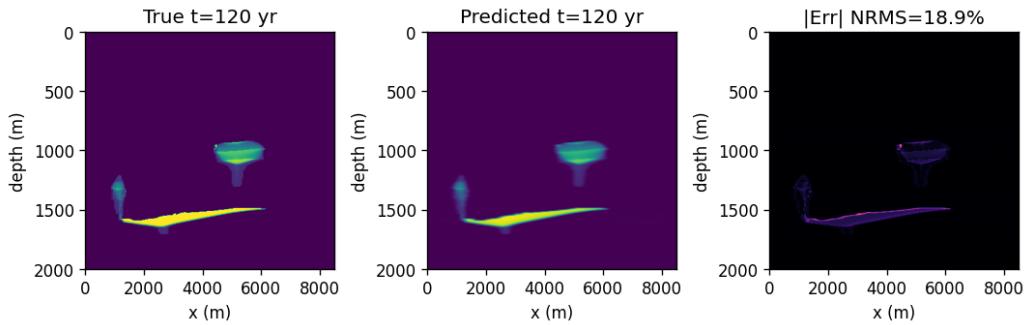


Figure 5.3: Representative frame at $t = 120$ yr. **Left:** ground truth ΔS . **Middle:** prediction. **Right:** absolute error with NRMS indicated.

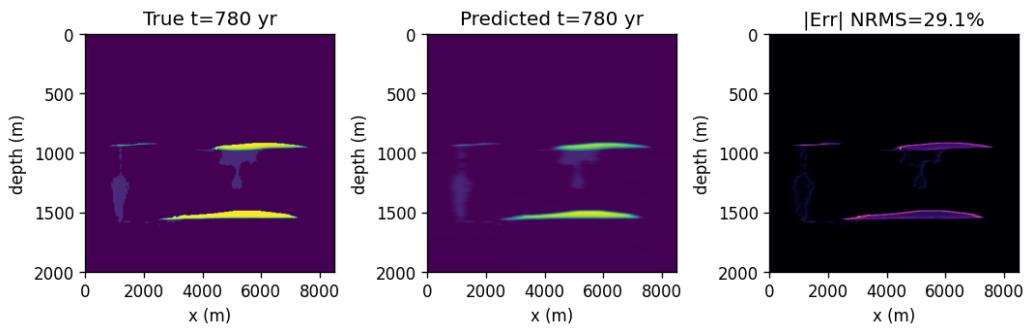


Figure 5.4: Representative frame at $t = 780$ yr. **Left:** ground truth ΔS . **Middle:** prediction. **Right:** absolute error with NRMS indicated.

Chapter 6

Conclusion

6.1 Summary

This thesis set out to recover time-lapse CO₂ saturation from angle gathers by combining a physics-guided forward operator with a learning model that exploits the structure of the data. The methodology (Chapter 2) constructs angle gathers by forward modeling from elastic logs, plane-wave (τ - p) decomposition, depth extrapolation with a smoothed kinematic background, and reparameterization to $\sin^2 \alpha$ for AVO consistency; per-angle wavelet shaping is estimated to reconcile ideal and measured amplitudes. Trained on 720 synthetic years and evaluated on 80 held-out years (Chapters 3-4), the network reproduces the large-scale kinematics of two migrating plumes across the full 800-year sequence. Plume locations, migration paths, and overall extents are consistently recovered in both seen and unseen intervals. To apply this framework to any specific case, it is important to make sure the data variability so that the generalization of the network can be guaranteed.

6.2 Main findings

The principal residuals are amplitude-related rather than geometric. Inside the plumes the predicted saturation exhibits a reduced dynamic range (attenuated gradients) and along plume rims the transitions are broadened. These effects are explained by the band-limit and aperture of the angle-gather imaging operator and by the smoothness-promoting regularization used during training; they do not displace the plumes. The NRMS history reflects three regimes: an elevated early transient while injection drives rapid geometric change; a long mid-sequence plateau with intermittent spikes coincident with abrupt structural updates (e.g., spill points and thin connectors); and

a modest step-up in the unseen interval that manifests as softer edges and slightly compressed contrasts while preserving correct localization. In short, the method is reliable for where the plumes are and how they move; caution is warranted when interpreting how sharp the fronts are or how strong interior contrasts appear.

Using a lightweight porosity mask as a soft spatial prior improves geometric fidelity by steering the model toward the reservoir architecture, but it introduces brittleness where the prior is violated (e.g., unmodelled leakage paths or faults). Treating the mask as a tunable, annealed prior and stress testing with deliberate perturbations provides a pragmatic compromise between guidance and adaptability. Because the training data are synthesized from known ground truth, uncertainty can be quantified directly; in field deployment, uncertainty should be communicated through ensembles (wavelet variability, background perturbations) and by reporting plume-extent and boundary metrics alongside NRMS.

6.3 Implications for monitoring

Given its strengths and biases, the model is well suited to tasks that prioritize plume localization and trajectory: early-warning screening, tracking lateral spread, and planning targeted follow-up measurements. The conservative amplitude bias is acceptable for these use cases and can be mitigated if quantitative saturation is required.

Overall, this work demonstrates that a physics-guided learning approach can recover the geometry and evolution of CO₂ plumes from angle gathers over centuries of simulated time. While amplitude sharpness remains the main limitation, the combination of principled physics, modest priors, and targeted learning offers a practical and extensible route toward robust, interpretable monitoring of subsurface storage.

References

- [1] F. M. Ahmed and D. B. S. Mohammed, “Feasibility of breast cancer detection through a convolutional neural network in mammographs,” *Tamjeed Journal of Healthcare Engineering Science and Technology*, vol. 1, no. 2, pp. 36–43, 2023.
- [2] K. Aki and P. G. Richards, *Quantitative Seismology: Theory and Methods*. New York: Freeman, 2002.
- [3] A. N. A. Ammah, “Applying time-lapse seismic inversion in reservoir management: A case study of the norne field,” Master’s thesis, Institutt for petroleumsteknologi og anvendt geofysikk, 2012.
- [4] M. Araya-Polo, S. Farris, and M. Florez, “Deep learning-driven velocity model building workflow,” *The Leading Edge*, vol. 38, no. 11, 872a1–872a9, 2019.
- [5] T. Bhakta, “Improvement of pressure-saturation changes estimations from time-lapse pp-avo data by using non-linear optimization method,” *Journal of Applied Geophysics*, vol. 155, pp. 1–12, 2018.
- [6] M. Bosch, T. Mukerji, and E. F. Gonzalez, “Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review,” *Geophysics*, vol. 75, no. 5, 75A165–75A176, 2010.
- [7] A. Buland and H. Omre, “Bayesian linearized AVO inversion,” *Geophysics*, vol. 68, no. 1, pp. 185–198, 2003.
- [8] W. Cao et al., “Accelerating 2d and 3d frequency-domain seismic wave modeling through interpolating frequency-domain wavefields by deep learning,” *Geophysics*, vol. 87, no. 4, T315–T328, 2022.
- [9] J. P. Castagna and H. W. Swan, “Principles of avo crossplotting,” *The Leading Edge*, vol. 16, no. 4, pp. 337–344, 1997.
- [10] A. Chadwick, G. Williams, N. Delepine, V. Clochard, K. Labat, S. Sturton, et al., “Quantitative analysis of time-lapse seismic monitoring data at the sleipner co₂ storage operation,” *The Leading Edge*, vol. 29, no. 2, pp. 170–177, 2010.

- [11] X. Chai, T. Yang, H. Gu, G. Tang, W. Cao, and Y. Wang, “Geophysics-steered self-supervised learning for deconvolution,” *Geophysical Journal International*, vol. 234, pp. 40–55, 2023.
- [12] C. Chen, Z. Chai, W. Shen, and W. Li, “Effects of impurities on CO_2 sequestration in saline aquifers: Perspective of interfacial tension and wettability,” *Industrial & Engineering Chemistry Research*, vol. 57, no. 1, pp. 371–379, 2018.
- [13] V. Das, A. Pollack, and U. Wollner, “Convolutional neural network for seismic impedance inversion,” *SEG Technical Program Expanded Abstracts*, 2018.
- [14] B. Dupuy, A. Romdhane, P. Eliasson, and H. Yan, “Combined geophysical and rock physics workflow for quantitative CO_2 monitoring,” *International Journal of Greenhouse Gas Control*, vol. 106, p. 103217, 2021.
- [15] J. L. Fatti, G. C. Smith, P. J. Vail, P. J. Strauss, and P. R. Levitt, “Detection of gas in sandstone reservoirs using avo analysis: A 3-d seismic case history using the geostack technique,” *Geophysics*, vol. 59, no. 9, pp. 1362–1376, 1994.
- [16] A.-K. Furre, O. Eiken, H. Alnes, J. N. Vevatne, and A. F. Kiær, “20 years of monitoring CO_2 injection at sleipner,” *Energy Procedia*, vol. 114, pp. 3916–3926, 2017. DOI: [10.1016/j.egypro.2017.03.1523](https://doi.org/10.1016/j.egypro.2017.03.1523).
- [17] J. Gao, Z. Li, and M. Zhang, “Seismic random noise attenuation based on m-resunet,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–16, 2023.
- [18] L. Gao et al., “Stransfuse: Fusing swin transformer and convolutional neural network for remote sensing image semantic segmentation,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021.
- [19] D. Grana, X. Lang, and W. Wu, “Statistical facies classification from multiple seismic attributes: Comparison between bayesian classification and expectation–maximization method and application in petrophysical inversion,” *Geophysical Prospecting*, vol. 65, no. 2, pp. 544–562, 2017.
- [20] D. Gray and E. Andersen, “The application of AVO and inversion to the estimation of rock properties,” in *SEG Technical Program Expanded Abstracts 2000*, Society of Exploration Geophysicists, 2000, SEG-2000-0549.

- [21] H. Hassani, A. Shahbazi, E. Shahbalayev, Z. Hamdi, S. Behjat, and M. Bataee, “Machine learning-based co₂ saturation tracking in saline aquifers using bottomhole pressure for ccs projects,” in *SPE Norway Subsurface Conference*, 2024.
- [22] G. Hema, S. P. Maurya, R. Kant, et al., “Seismic inversion for co₂ volume monitoring and comprehensive evaluation of pore fluid properties: A case study,” *Environmental Earth Sciences*, vol. 84, p. 72, 2025. DOI: 10.1007/s12665-025-12088-5.
- [23] D. Hendrycks and K. Gimpel, *Gaussian error linear units (GELUs)*, 2016. arXiv: 1606.08415. [Online]. Available: <https://arxiv.org/abs/1606.08415>.
- [24] Q. Hu, D. Grana, and K. A. Innanen, “Feasibility of seismic time-lapse monitoring of co₂ with rock physics parametrized full waveform inversion,” *Geophysical Journal International*, vol. 233, 2023. DOI: 10.1093/gji/ggac462.
- [25] V. Kazei, A. Titov, W. Li, et al., “Predicting density and velocity ahead of the bit with zero-offset vsp using deep learning,” in *First International Meeting for Applied Geoscience & Energy (IMAGE)*, SEG, 2021, pp. 1450–1454.
- [26] S. Khandoozi, R. Hazlett, and M. Fustic, “A critical review of co₂ mineral trapping in sedimentary reservoirs—from theory to application: Pertinent parameters, acceleration methods and evaluation workflow,” *Earth-Science Reviews*, vol. 244, p. 104515, 2023.
- [27] S. Kim, J. Park, S. J. Seol, and J. Byun, “Machine learning-based time-lapse 1d seismic full-waveform inversion with efficient training data generation in a carbon capture and storage monitoring,” *Geoenergy Science and Engineering*, vol. 238, p. 212852, 2024.
- [28] K. Kitamura et al., “The potential of v_s and v_p–v_s relation for the monitoring of the change of co₂-saturation in porous sandstone,” *International Journal of Greenhouse Gas Control*, vol. 25, pp. 54–61, 2014.
- [29] M. Landrø and M. Zumberge, “Estimating saturation and density changes caused by co₂ injection at sleipner—using time-lapse seismic amplitude-variation-with-offset and time-lapse gravity,” *Interpretation*, vol. 5, no. 2, T243–T257, 2017.
- [30] Y. LeCun and Y. Bengio, “Convolutional networks for images, speech, and time series,” in *The Handbook of Brain Theory and Neural Networks*, 1995, pp. 3361–3366.

- [31] Y. Liu and Y. Zhong, “Machine learning-based seafloor seismic prestack inversion,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4471–4480, 2020.
- [32] P. J. Loftus, A. M. Cohen, J. C. Long, and J. D. Jenkins, “A critical review of global decarbonization scenarios: What do they tell us about feasibility?” *Wiley Interdisciplinary Reviews: Climate Change*, vol. 6, no. 1, pp. 93–112, 2015.
- [33] S. Mallick and S. Adhikari, “Amplitude-variation-with-offset and prestack-waveform inversion: A direct comparison using a real data example from the rock springs uplift, wyoming, USA,” *Geophysics*, vol. 80, no. 2, B45–B59, 2015.
- [34] S. P. Maurya, N. P. Singh, and K. H. Singh, *Seismic inversion methods: a practical approach*. Springer, 2020, vol. 1, p. 694.
- [35] A. P. Muller et al., “Deep-tomography: Iterative velocity model building with deep learning,” *Geophysical Journal International*, vol. 232, no. 2, pp. 975–989, 2023.
- [36] P. Newell and A. G. Ilgen, *Science of Carbon Storage in Deep Saline Formations*. Elsevier, 2019, pp. 1–13. DOI: [10 . 1016 / B978 - 0 - 12 - 812752 - 0 . 00001 - 0](https://doi.org/10.1016/B978-0-12-812752-0.00001-0).
- [37] L. Niu, J. Geng, X. Wu, L. Zhao, and H. Zhang, “Data-driven method for an improved linearised avo inversion,” *Geophysical and Geological Engineering*, vol. 18, 2020.
- [38] J. M. Nordbotten, M. A. Fernø, B. Flemisch, A. R. Kovscek, and K.-A. Lie, “The 11th society of petroleum engineers comparative solution project: Problem definition,” *Society of Petroleum Engineers Journal*, 2024.
- [39] E. B. Raknes, B. Arntsen, and W. Weibull, “Three-dimensional elastic full waveform inversion using seismic data from the sleipner area,” *Geophysical Journal International*, vol. 202, no. 3, pp. 1877–1894, 2015.
- [40] Y. Ren, X. Xu, S. Yang, L. Nie, and Y. Chen, “A physics-based neural-network way to perform seismic full waveform inversion,” *IEEE Access*, vol. 8, pp. 112 266–112 277, 2020.
- [41] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *MICCAI – Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234–241.

- [42] B. H. Russell, “Prestack seismic amplitude analysis: An integrated overview,” *Interpretation*, vol. 2, no. 2, SC19–SC36, 2014.
- [43] B. H. Russell, D. Gray, and D. P. Hampson, “Linearized avo and poroelasticity,” *Geophysics*, vol. 76, no. 3, pp. C19–C29, 2011.
- [44] R. T. Shuey, “A simplification of the zoeppritz equations,” *Geophysics*, vol. 50, no. 4, pp. 609–614, 1985.
- [45] J. Simon, G. Fabien-Ouellet, E. Gloaguen, and I. Khurjekar, “Hierarchical transfer learning for deep learning velocity model building,” *Geophysics*, vol. 88, no. 1, R79–R93, 2023.
- [46] L. Sirgue, O. I. Barkved, J. P. Van Gestel, O. J. Askim, and J. H. Kommedal, “3d waveform inversion on valhall wide-azimuth obc,” in *71st EAGE Conference and Exhibition incorporating SPE EUROPEC 2009*, 2009.
- [47] F. Tamburini, F. Zanobetti, M. Cipolletta, S. Bonvicini, and V. Cozzani, “State of the art in the quantitative risk assessment of the ccs value chain,” *Process Safety and Environmental Protection*, vol. 191, pp. 2044–2063, 2024. doi: [10.1016/j.psep.2024.09.066](https://doi.org/10.1016/j.psep.2024.09.066).
- [48] E. S. Um, D. Alumbaugh, M. Commer, et al., “Deep learning multi-physics network for imaging co₂ saturation and estimating uncertainty in geological carbon storage,” *Geophysical Prospecting*, vol. 72, no. 1, pp. 183–198, 2023.
- [49] D. Voskov et al., *Open delft advanced research terra simulator (open-darts)* (v1.2.2), 2019. doi: [10.5281/zenodo.14747732](https://doi.org/10.5281/zenodo.14747732).
- [50] L. Wei, P. Roy, T. Dygert, D. Grimes, and M. Edwards, “Estimation of reservoir pressure and saturation changes from 4d inverted elastic properties,” *SEG Technical Program Expanded Abstracts*, pp. 5923–5927, 2017.
- [51] J. W. Wiggins, “Attenuation of complex water-bottom multiples by wave-equation-based prediction and subtraction,” *Geophysics*, vol. 53, no. 12, pp. 1527–1539, 1998.
- [52] Z. Yin, H. T. Erdinc, A. P. Gahlot, M. Louboutin, and F. J. Herrmann, “Derisking geologic carbon storage from high-resolution time-lapse seismic to explainable leakage detection,” *The Leading Edge*, 2024. doi: [10.1190/tle42010069.1](https://doi.org/10.1190/tle42010069.1).
- [53] Z. Yin, M. Louboutin, O. Møyner, and F. J. Herrmann, “Time-lapse full-waveform permeability inversion: A feasibility study,” *The Leading Edge*, 2024. doi: [10.1190/tle43080544.1](https://doi.org/10.1190/tle43080544.1).

- [54] H. Zhang et al., “Automatic velocity analysis using interpretable multimode neural networks,” *Geophysical Journal International*, vol. 235, no. 1, pp. 216–230, 2023.
- [55] L. Zhao et al., “Fluid and lithofacies prediction based on integration of well-log data and seismic inversion: A machine-learning approach,” *Geophysics*, vol. 86, no. 4, pp. M151–M165, 2021.
- [56] Y. Zheng, “Elastic pre-stack seismic inversion in stratified media using machine learning,” in *81st EAGE Conference and Exhibition*, 2019.
- [57] Z. Zhong and T. R. Carr, “Geostatistical 3d geological model construction to estimate the capacity of commercial scale injection and storage of co₂ in jacksonburg-stringtown oil field, west virginia, usa,” *International Journal of Greenhouse Gas Control*, vol. 80, pp. 61–75, 2019.
- [58] Z. Zhong, A. Y. Sun, Q. Yang, and Q. Ouyang, “A deep learning approach to anomaly detection in geological carbon sequestration sites using pressure measurements,” *Journal of Hydrology*, vol. 573, pp. 885–894, 2019.
- [59] K. Zoeppritz, “Erdbebenwellen viii b, über reflexion und durchgang seismischer wellen durch unstetigkeitsflächen,” *Göttinger Nachr*, vol. 1, pp. 66–84, 1919.

Appendix

Network hyperparameters

Table 6.1: Training hyperparameters.

Parameter	Value
Optimizer	Adam
Learning rate	1×10^{-4}
Batch size	8
Loss	Eq. (3.8)
Training epochs	60
Angle set	$\sin^2 \alpha \in \{0.0, 0.1, 0.2, 0.3\}$

Extra figures

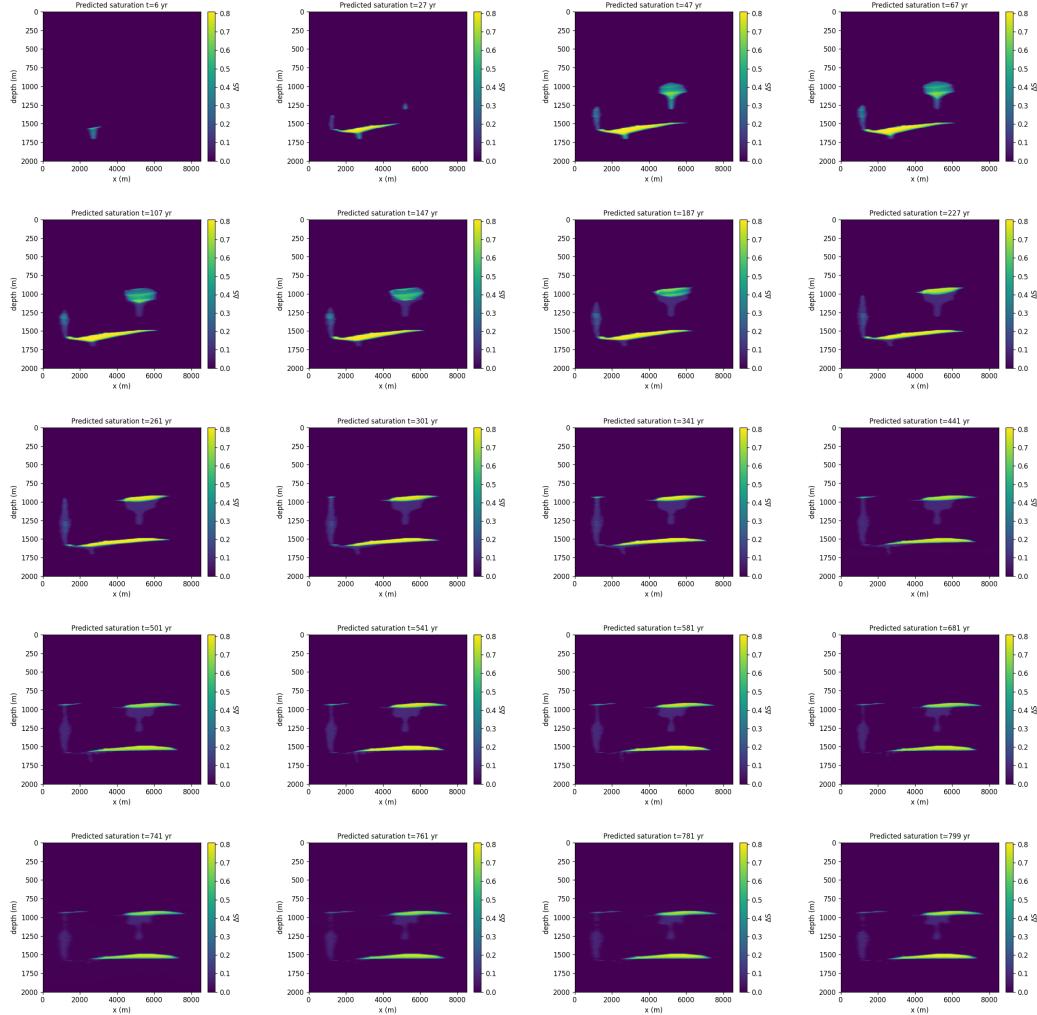


Figure 6.1: Sampled snapshots of predicted CO₂ migration over 800 years.

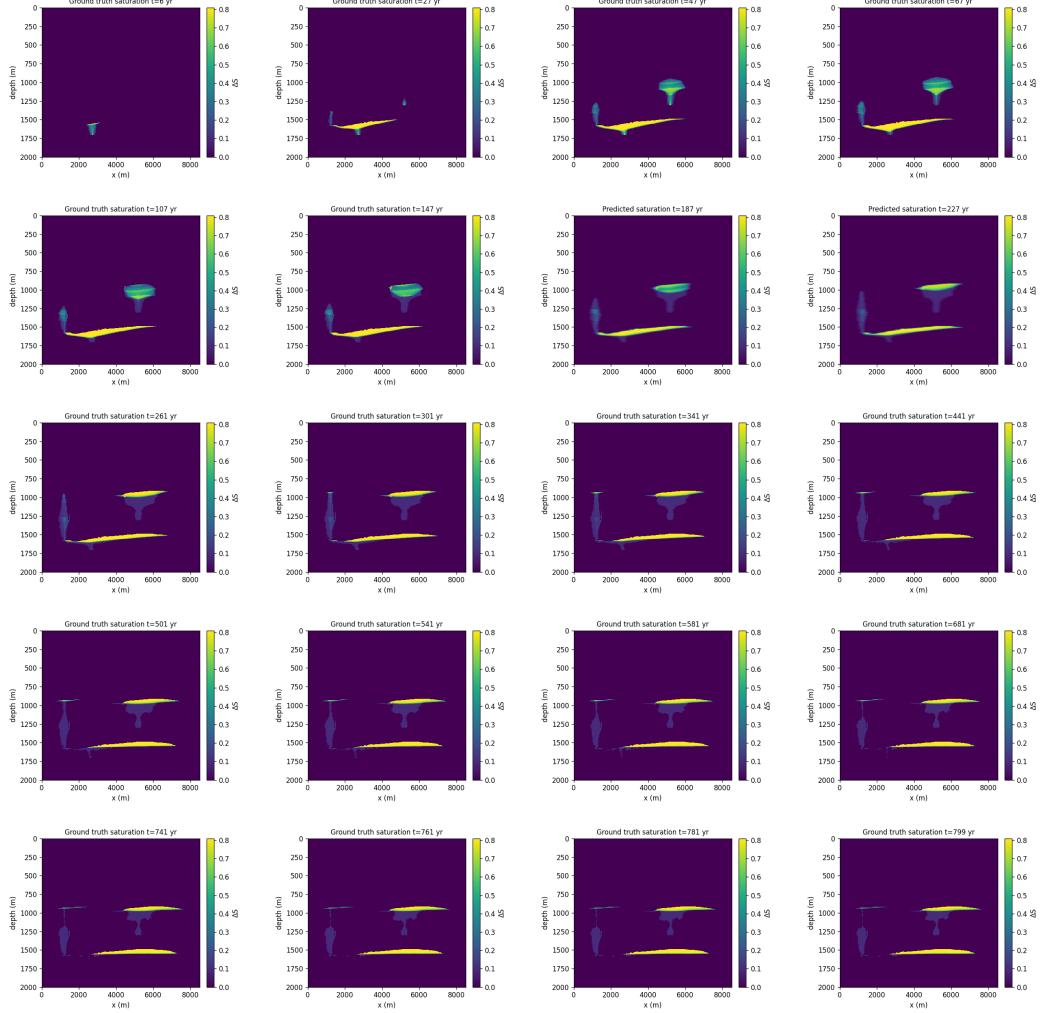


Figure 6.2: Sampled snapshots of ground-truth CO_2 migration over 800 years.