A Strategy for Using Machine Learning in Modelling the Seismic Response to CO_2 Storage

[1,2]Adewunmi Abdul*, [1]Tarje Nissen-Meyer and [1]J-Michael Kendall
[1]University of Oxford, United Kingdom
[2]Petroleum Development Trust Fund (PTDF), Nigeria

ABSTRACT

Finite difference numerical modelling is a common technique for simulating seismic waves and can be used for detailed characterization of the seismic response to CO_2 migration within reservoir units. These large synthetic datasets also lend themselves well to machine-learning techniques, both for modelling and analyses. Our research aims to combine the knowledge acquired from numerical modelling of wave propagation with machine learning to better understand how migrates within reservoir units over long periods of time at computationally reasonable costs. We outline a finite difference methodology that simulates waveforms for a large range of synthetic structural scenarios based on the properties of the clastic rocks and fluids from the Utsira formation in the Sleipner field, North Sea. Transfer learning from a previously developed model, which was used to generate the seismic response of faulted media is employed in our neural network to simulate the waveform signature of sequestered CO_2 in our models. Although when presented with models outside the range of our training distribution, the network's accuracy reduces. In the future we discuss analysing the sensitivity of CO_2 migration to seismic waveforms to help us understand the sensitivity of seismic waves to realistic storage scenarios.

Keywords: Machine learning, Numerical modelling, migration, Carbon dioxide storage, Seismic response

INTRODUCTION

Rapidly increasing global population leads to phenomenal energy demand, not only for lighting our homes, but also for transportation and to power several industrial applications to produce goods and services for the comfort and efficiency of life. Today, more than 80% of global energy demand is met by fossil fuels. Nonetheless, fossil fuels are the leading cause of global warming, and the biggest environmental threat to the liveability of our planet. Because fossil fuel is mainly composed of carbon, its combustion releases substantial amounts of carbon dioxide - a leading player in global greenhouse gas emissions.

In 2021, global carbon dioxide *(CO_2)* emissions resulting from energy combustion and industrial activities experienced a notable rebound, reaching their peak annual level. This surge, which saw a 6% rise compared to 2020, brought emissions to an estimated 36.3 Gt (IEA, 2022). Although renewable energy options appear promising for a net-zero emission future, fossil fuels remain a key player in the energy mix, at least regionally and temporarily. This fact highlights the critical need for efficient means to capture, store and utilize global CO_2 emissions. Just as how fossil fuels (such as oil and gas) are produced from the subsurface, geological CO_2 storage is conceptually as simple. Captured CO_2 can be injected back into the geological subsurface in places such as depleted oil and gas reservoirs, deep saline aquifers and unmineable coal beds. The success of such storage, however, requires effective monitoring methods to ensure the stored CO_2 is appropriately contained in the subsurface without any risk of leakage. One popular approach of such leakage and containment monitoring relies on geophysical observations and monitoring sensor networks.

Of the various CO_2 monitoring techniques, repeated seismic surveys (also known as 4D seismic or time-lapse seismic) remains the most common method for tracking the subsurface migration of stored CO_2 (Figure 1) over time. Time-lapse seismic has the ability to provide constraints.

The authors thank the Petroleum Technology Development Fund (PTDF) of Nigeria for the scholarship. We also thank Ben Moseley for his fast-seismic simulation in 2-D faulted acoustic media code, which is available on GitHub (https://github.com/benmoseley/seismic-simulation complex-media). We appreciate the Computational Infrastructure for Geodynamics (https://www.geodynamics.org/) for availing us the open-source SEISMIC_CPML FD modelling libraries and the Advanced Research Computing (ARC), University of Oxford for providing access to High Performance Computing (HPC) resources.

needed to detect changing rock property related to changes in fluid, pressure and temperature conditions in the host rock formations (IPCC, 2005, The Royal Society, 2022). Such constraints are powerful as they potentially improve our understanding and visualization of the progressive development of CO_2 plumes or migration pathways (Chadwick *et al,* 2009, Chadwick *et al,* 2010, Chadwick and Noy, 2010).

However, seismic surveys can be expensive to conduct as they involve special equipment, acquisition, processing, and interpretation. The cost of regularly conducting seismic surveys may not be justified, especially for long-term CO_2 monitoring where ongoing data collection and interpretation can quickly become prohibitively expensive. The use of numerical methods, governed by the discretization of the wave equation allow for the generation of enough synthetic data for proper understanding and modelling of CO_2 plume development (Arts *et al,* 2003, Carcione *et al,* 2006, Rubino *et al,* 2011, Williams and Chadwick, 2021). The application of neural network to these methods, can provide a significant increase in computational speed as seismic responses are output in one inference step rather than iteratively modelling the seismic wavefield through time (Moseley *et al,* 2020). In addition, neural network helps identify previously unknown links between input data and output to make predictions in unseen input examples (Mattéo *et al.,* 2021) in order to give accurate, consistent and objective output models (Li and Li, 2021).

METHODOLOGY

Building models

We build 2D models based on the Sleipner field in the North Sea, where CO_2 is currently been stored in its Utsira formation. Models are 2048m in width and depth and we use the rock and fluid data gathered from the field area (Table 1). The model is comprised of the water column, overburden, caprock, reservoir and underlayer or underburden, assuming scenarios for before and after CO_2 injection. To start, the reservoir rock contains brine. CO_2 is then added in varying saturation percentages, thicknesses and column widths after injection. The generated models were carefully visualized to check for unrealistic models. Example models are shown in Figure 2.

Table 1: Rock and fluid data gathered from the Sleipner field and used to construct seismic models. The range in parameter values considered in model building is shown

Reservoir properties	Range
P wave velocity: brine sand (m/s)	1950-2100
Density: brine sand (Kg/m^3)	2000-2150
Reservoir thickness (m)	20-300
Caprock thickness (m)	6-150
CO_2 saturation (%)	0 - 90
CO_2 plume thickness (m)	0- 70
CO_2 plume width (m)	0 – 400

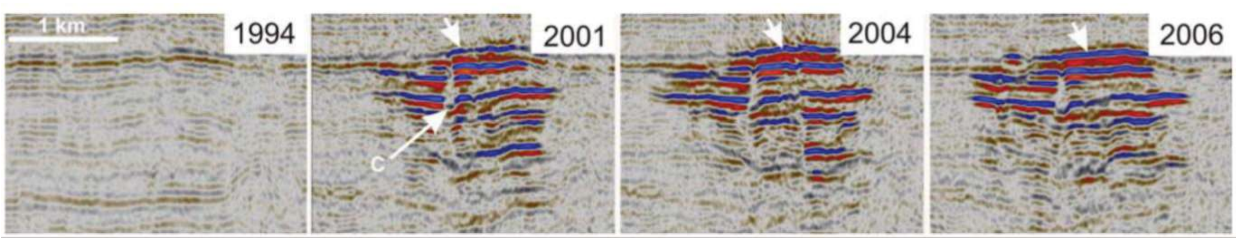

Figure 1: Time-lapse seismic images illustrating the CO_2 plume development, spanning the 1994 dataset (before injection) to the 2006 datasets (after injection) (Adapted from Chadwick and Noy, 2010).

Motivation

Here we consider the use of finite difference numerical modelling method for simulating the seismic response of CO_2 storage in saline aquifer. Our aim is to simulate and predict the seismic signature of CO_2 storage over a long period of time, which is crucial for the success of CCS. Additionally, we apply machine learning algorithms to these numerical model examples, to uncover hidden patterns and correlations between the model examples and their pressure responses at a reduced computational time. This thereby enables us to make data-driven decisions and gain deeper insights into subsurface CO_2 behaviour.

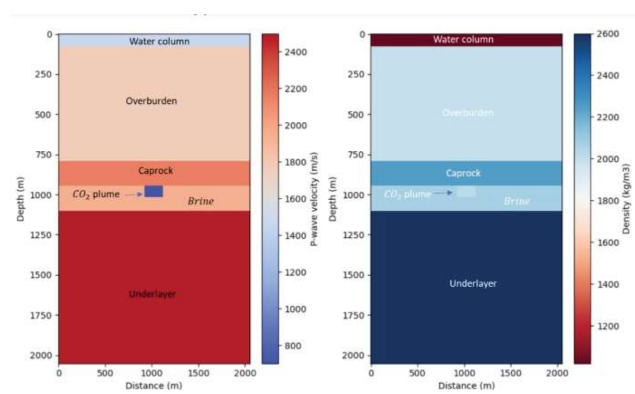

Figure 2: An example of a velocity (left) and density (right) model where 20% CO_2 is injected into a lens 70 m in thickness and 200.0 m in width.

FD simulations

Our 2D earth model is divided into 256 by 256 grid points, with an 8m separation. We use a 10 Hz dominant frequency Ricker source at a depth of 16 m in water and record the corresponding pressure response at 128 receiver locations, each placed 15 m apart on the seabed. A time-stepping scheme is used and implementing Komatitsch and Martin (2007)'s SEISMIC_CPML solver, which uses the finite difference method to solve the discretized equations over time. The numerical solution calculates pressure changes at each grid point during each time step. Simulation results are collected at different time steps to visually validate how the pressure field evolves over time.

CNN strategy

We consider a combination of 30,000 input material property (velocity and density) models and their corresponding pressure responses. These serve as ground truth for the neural network models. To accomplish this, we build on Moseley et al (2020)'s pre-trained neural network models, which successfully used the encoder-decoder architecture to learn how to simulate seismic waves in 2-D faulted acoustic media. Here we apply this approach to modelling the seismic response of CO_2 storage.

With two model inputs in our case, and a larger grid size, we add additional layers to both the encoder and decoder to accommodate our larger input sizes. The encoder takes the two-input data and compress them into a lower-dimensional representation called the "latent space." The decoder on the other hand, takes this latent space representation and generates synthetic seismic wavefields, matching the target data (Figure 3).

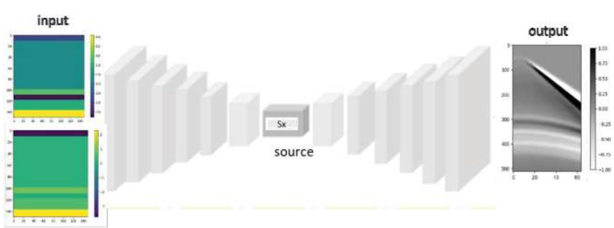

Figure 3: Encoder-Decoder Neural Network Architecture (Adapted from Li *et al.,* 2020). Given material property inputs, source location added to the latent vector, the network outputs the pressure responses at the receiver locations.

Training Process:

We train the adapted architecture on our generated synthetic dataset. The dataset is split into training and validation sets in the ratio 24,000:6,000 for training the model and monitoring its performance during training, respectively. We fine-tune the model's weights and biases to optimize it for the task. The training is carried out in batches of 50 to speed up convergence and reduce memory requirements. The model learns to generate synthetic seismic wavefields that match the target data. We continuously monitor the model's performance on the validation set during training by checking for overfitting. Once training is complete, we evaluate the model on 3000 unseen test dataset to assess the model's capabilities.

RESULTS AND DISCUSSION

We create 13,365 different velocity models representing different cases of before and after Co_2 injection. The velocity and density values of the rock and fluid are consistent with plausible geological expectations (Figure 2). Our FD solver is able to generate the pressure response at different receiver positions for 10,000 velocity and density models, each with 3 random source locations, thereby resulting in 30,000 simulations (Figure 4). The encoder-decoder NN accurately captures the patterns between the input data and the label, thereby giving outputs that correctly simulate the receiver response of the wave. The network is able to generalize to new, unseen data (Figure 5) within the range of training data.

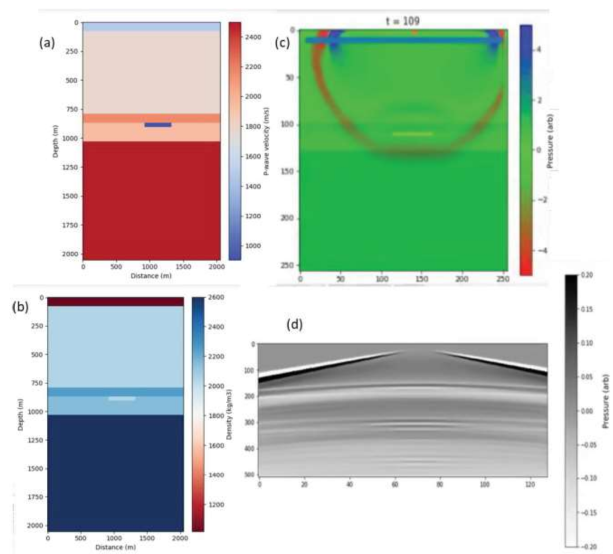

Figure 4: FD simulation example for a 400 m wide and 35 m thick plume of 2% 2. (a) Velocity model; (b) Density model; (c) Wavefield for a moment in time (the source location is shown as a red dot, the receivers lie along the horizontal blue line below the source); (d) An example of the pressure response recorded at the receiver locations.

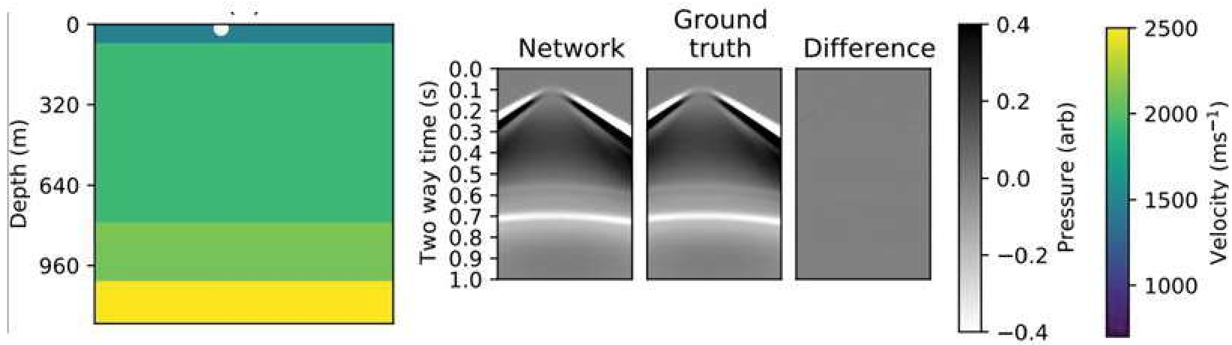

Figure 5: Unseen test dataset example simulation. From left to right: Velocity model with source location (white circle); network prediction; FD simulation; and the difference between the two.

CONCLUSION

The seismic method is a good CO_2 monitoring tool; we build models to help us understand how waveforms behave with respect to changes in the subsurface for early remediation. Neural network can simulate the seismic response to changes in CO_2 concentration and location in simple generic models. The network performs well on new unseen data but struggles once the unseen data are for models outside the range of those in the training data frame. Our next step is to systematically explore models of CO_2 storage, assessing how well seismic reflection surveys can detect differences between models. This sensitivity analysis is an important step in interpreting time-lapse seismic data.

REFERENCES CITED

Arts, R., Eiken, O., Chadwick, R. A., Zweigel, P., van Der Meer, L. & Zinszner, B. (2003). Monitoring of CO_2 injected at Sleipner using time lapse seismic data.In: Gale, J. & Kaya, Y. (eds) Greenhouse Gas Control Technologies. Elsevier Science, Oxford, 347–52.

Carcione, J., Picotti, S., Gei, D., & Rossi, G. (2006). Physics and seismic modelling of monitoring CO_2 storage, Pure Appl. Geophys.,163, 175–207.

Chadwick, R.A., Noy, D., Arts, R., & Eiken O. (2009). Latest time-lapse seismic data from Sleipner yield new insights into CO_2 plume development, Energy Procedia, Volume 1, Issue 1, ISSN 1876-6102, https://doi.org/10.1016/j.egypro.2009.01.274, 2103-2110.

Chadwick, A., Gareth, W., Nicolas, D., Vincent, C., Karine, L., Susan, S., Maike, B., Menno, D., Michael, N., Anne, L.L., Rob, A., Filip, N., & Giuliana, R. (2010). Quantitative analysis of time-lapse seismic monitoring data at the Sleipner CO_2 storage operation, The Leading Edge 29:2, 170-177.

Chadwick, R.A. & Noy, D.J. (2010). History – matching flow simulations and time-lapse seismic data from the Sleipner CO_2 plume. In: Vining, B.A. and Pickering, S.C. (Eds.) Petroleum Geology: From Mature basins to New Frontiers – Proceedings of the 7th Petroleum Geology Conference, Geological Society, London, 1171-1182.

IEA, (2022): Global Energy Review: CO_2 Emissions in 2021 Global emissions rebound sharply to highest ever level.

IPCC, (2005): IPCC Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change [Metz, B., O. Davidson, H. C. de Coninck, M. Loos, & L. A. Meyer (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 442.

Komatitsch, D. & Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation, Geophys. J. Int., 139, 806–822.

Li, B., & Li, Y. E. (2021). Neural network-based CO_2 interpretation from 4D Sleipner seismic images. Journal of Geophysical Research: Solid Earth, 126, e2021JB022524. https://doi.org/10.1029/2021JB022524.

Li, J., li, G., & Fan, H. (2020). Image Reflection Removal Using End-to-End CNN. IET Image Processing. doi:10.1049/iet-ipr.2019.0247.

Mattéo, L., Manighetti, I., Tarabalka, Y., Gaucel, J.-M., Van den Ende, M., Mercier, A., et al. (2021). Automatic fault mapping in remote optical images and topographic data with deep learning. Journal of Geophysical Research, Solid Earth, 126, e2020JB021269. https://doi.org/10.1029/2020JB021269.

Moseley, B., Nissen-Meyer, T., & Markham, A. (2020). Deep learning for fast simulation of seismic waves in complex media, Solid Earth, 11, 1527–1549, https://doi.org/10.5194/se-11-1527-2020.

Rubino, J. G., Velis D. R., & Sacchi, M. D. (2011). Numerical analysis of wave-induced fluid flow effects on seismic data: Application to monitoring of CO_2 storage at the Sleipner field, J. Geophys. Res.,116, B03306, doi:10.1029/2010JB007997.

The Royal Society. (2022). Locked away – Geological carbon storage policy briefing. ISBN: 978-1-78252-616-2.

Williams G.A. and Chadwick R.A. (2021). Influence of reservoir-scale heterogeneities on the growth, evolution and migration of a CO_2 plume at the Sleipner Field, Norwegian North Sea, International Journal of Greenhouse Gas Control. Volume 106, 103260, ISSN 1750-5836, https://doi.org/10.1016/j.ijggc.2021.103260.