CCUS: 4186155

Estimating CO_2 Saturation in Subsurface Geologic Reservoirs Using Pre-stack Seismic Attributes and Machine Learning

Papa Owusu*[1], Rui Zhang[1], 1. University of Louisiana at Lafayette.

This paper was prepared for presentation at the Carbon Capture, Utilization, and Storage conference held in Houston, TX, 03-05 March.

Abstract

Integrating well-logs and time-lapse seismic data to estimate spatial CO_2 saturation in reservoirs is essential to geological carbon storage, monitoring, and management. This study leveraged pre-stack Amplitude-vs-Angle (AVA) attributes with machine-learning techniques to estimate CO_2 saturation in the lower Tuscaloosa formation, Cranfield site, Mississippi, USA. Random Forest (RF) and XGBoost (XGB) regression models trained yielded low mean absolute errors of 0.029 and 0.028, respectively, on synthetic blind test data. Notwithstanding, only the RF model predicted geologically plausible CO_2 saturations from the field time-lapse pre-stack data. Inference from the CO_2 saturation distribution in the anticline reservoir indicates slow migration of the injected supercritical carbon from the injection zone down-dip to the crest of the structure.

Introduction

Seismic surveys remain critical in addressing the three main monitoring objectives of carbon storage in geologic reservoirs: conformance, containment, and contingency (Ringrose, 2023). In recent years, the integration of post-stack seismic data forms and machine learning algorithms have been utilized to characterize CO_2 plumes and estimate saturations in geologic reservoirs (Hussein et al., 2021; Leong et al., 2024; Li et al., 2021; Xue et al., 2023). While pre-stack seismic attributes have been instrumental in conventional hydrocarbon exploration and production management, their application in CO_2 saturation estimation remains underexplored. This study leverages pre-stack attributes as features for supervised machine learning regression algorithms to estimate CO_2 saturation.

The study focuses on the Cranfield site in Mississippi, USA. Over 2 million metric tonnes of CO2 have been injected into the anticlinal lower Tuscaloosa formation at depths between 3012 and 3142 meters (Zhang et al., 2013a).

Methods

Our workflow encompasses three steps: synthetic data generation, machine learning (ML) modeling, and field data inference. Baseline well-log data measuring porosity, bulk density, and P-wave and S-wave velocities were sampled at the reservoir columns of four wells penetrating the lower Tuscaloosa formation: injector well F1 and monitoring wells F2, F3, and W28. Each well was sampled at 1ft intervals within the reservoir columns for fluid substitution, post-stack attribute generation, and subsequent compiling as labeled data for machine learning training and validation. The Gassmann fluid substitution approach with homogeneous saturation was used to simulate new densities and velocities at an incremental 10% change in CO_2 saturation from the baseline case (0% CO_2 saturation). The average reservoir mineral composition ratio used in fluid substitution computation for this study is 79.4% quartz, 11.8% chlorite, 3.1% kaolinite, 1.3% illite, 1.5% calcite with dolomite, and 0.2% feldspar (Lu et al., 2012). The elastic properties for the baseline and each CO_2 substituted case were used to compute pre-stack attributes: intercept, gradient, lambda-rho, mu-rho, Poisson's ratio, and P-wave impedance. The intercept and gradient attributes were calculated using Shuey's approximations at Amplitude-vs-Angle (AVA) stacks ranging between 0° and 30° with steps of 3°. The lambda-rho and mu-rho attributes are the product of incompressibility with bulk density and rigidity with bulk density, respectively. The calculated P-wave acoustic impedance attribute was also normalized to 0.5*In(P-wave impedance). A labeled tabular data was compiled with CO_2 saturation as the target (dependent variable) and each attribute difference (baseline minus fluid-substituted case) for every sampled well data point as features (independent variables) for ML modeling. The attributes were computed using the open-source Python library Bruges.

Random Forest (RF) and XGBoost (XGB) regression models were built using the Sci-kit Learn Python library for machine learning training and validation for this project. The hyperparameters selected for the final RF model include number of trees (n_estimators = 66), maximum depth of each tree (max_depth = 12), and maximum number of leaf nodes per tree (max_leaf_nodes = 200). The final XGB model has hyperparameters, number of trees (n_estimators =120), maximum depth of a tree (max_depth = 4), maximum number of leaves per tree (max_leaves = 15), and a minimum child weight (min_child_weight = 8.1). An 80-20% train-test data split and a further 50-50% split of the test dataset into validation and blind test samples were implemented to build models with high accuracy and optimum regularization. Pre- and post-injection pre-stack seismic data volumes of the Cranfield field were acquired in 2007 and 2010, respectively. The time-lapse field seismic volumes, registered by Zhang et al. (2013), were used for the pre-stack attribute inversion. Each attribute difference (baseline minus post-injection) was extracted as an interval 5ms above the top and below the base reservoir horizons and compiled as features for CO_2 saturation estimation.

Results and Discussion

Both tuned RF and XGB models yielded training scores (r^2) of 0.99, whereas validation scores were 0.97 and 0.98, respectively. Blind test evaluation on both models yielded mean absolute errors of 0.029 and 0.028, respectively. Figures 1 and 2 highlight the CO_2 saturation distribution estimated by the RF and XGB models. The section is crossline 197, which parallels the trend of the reservoir anticline structure and intersects the wells F1 (CO_2 injector), F2 (monitoring), and F3 (monitoring) down-dip. Comparatively, the fluid distribution signatures of both models in the reservoir are consistent. However, the CO_2 saturation estimation by the RF model ranges from 0 to 100%, whereas that of the XGB model ranges from -30% to 120%, which is implausible. Figures 3 and 4 highlight CO_2 plumes within the reservoir, near the apex (time slice 2250 ms) and down-dip (time slice 2284 ms). It is observed that high CO_2 saturations are proximal to the injection zone of the reservoir relative to the crest.

Despite the robustness of both models in terms of accuracy and generalization on the synthetic well data, the XGBoost model underperforms in estimating CO_2 saturation from field seismic data. This could be attributed to its inability to map learned weights from the high-resolution well logs to low-resolution seismic data. Assessment of the CO_2 plume at various elevations within the reservoir structure indicates slow migration of the injected supercritical carbon from the injection zone down-dip to the crest of the structure. The gentle dip of the reservoir structure, coupled with reservoir heterogeneities, could impede CO_2 migration. Another contributory factor could be the short time-lapse interval between the baseline pre-injection (2007) and monitor post-injection (2010) seismic survey. Therefore, later repeat surveys could ascertain the extent of the spatial distribution of CO_2 saturation within the reservoir structure.

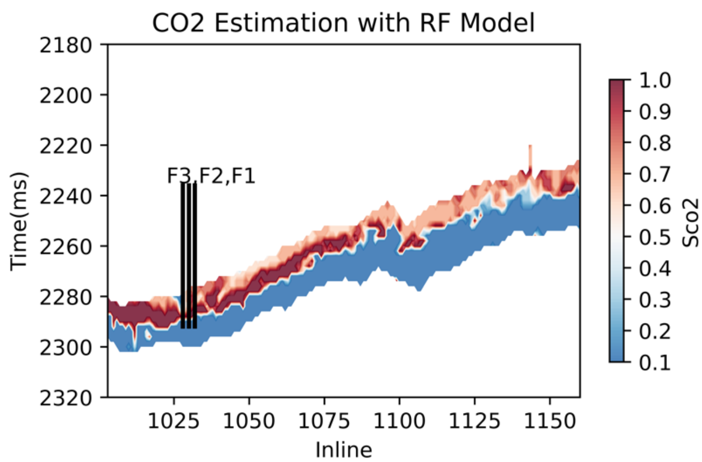

Figure 1. Crossline 197 shows CO_2 saturation distribution in the reservoir estimated by the Random Forest (RF) model. The black lines mark the location of wells F1, F2, and F3.

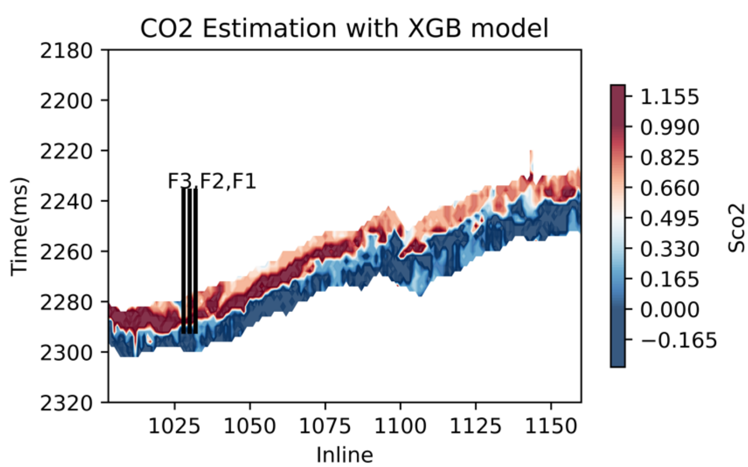

Figure 2. Crossline 197 shows CO_2 saturation distribution in the reservoir estimated by the XGBoost (XGB) model. The black lines mark the location of wells F1, F2, and F3.

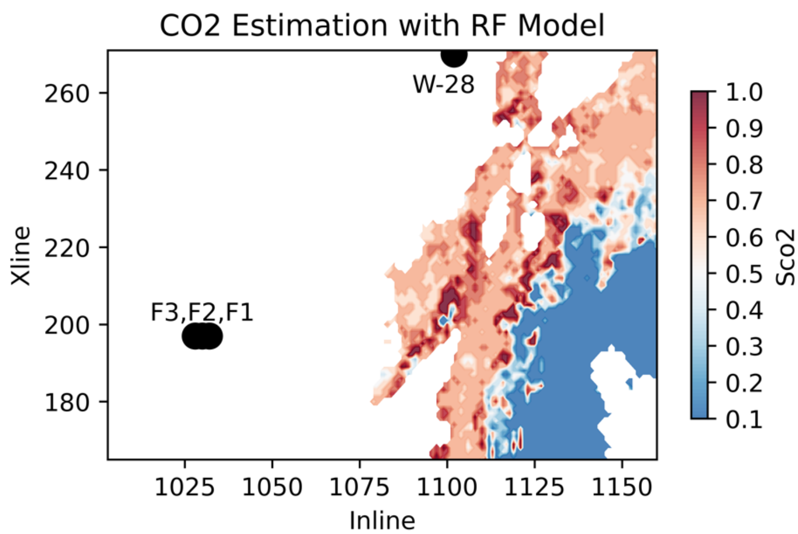

Figure 3. Time slice at 2250 milliseconds, highlighting CO_2 saturation distribution up-dip of the reservoir structure. The black dots mark well locations F1, F2, F3, and W-28.

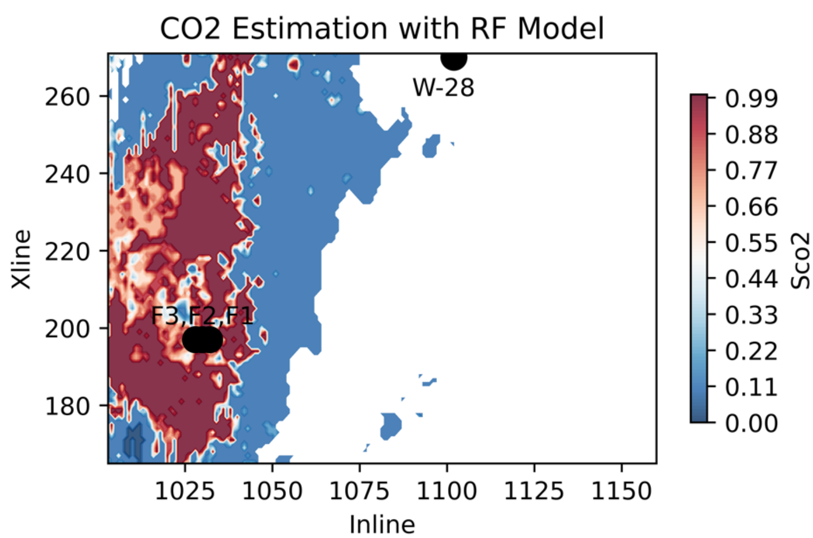

Figure 4. Time slice at 2284 milliseconds, highlighting CO_2 saturation distribution down-dip of the reservoir structure. The black dots mark well locations F1, F2, F3, and W-28.

Conclusions

The study demonstrates the sensitivity of pre-stack seismic attributes and machine learning to estimate CO_2 saturation. This approach provides a valuable tool for field-scale carbon sequestration monitoring and management. Further, thin-bed reservoir effects on seismic amplitude signatures, such as interference and tuning, must be investigated and incorporated to build robust models sensitive to field scale complexities.

References

Hussein, M., Stewart, R. R., Sacrey, D., Johnston, D. H., & Wu, J. (2021). Unsupervised machine learning for time-lapse seismic studies and reservoir monitoring. *Interpretation*, *9*(3), T791–T807. https://doi.org/10.1190/INT-2020-0176.1

Leong, Z. X., Zhu, T., & Sun, A. Y. (2024). Time-lapse seismic inversion for CO2 saturation with SeisCO2Net: An application to Frio-II site. *International Journal of Greenhouse Gas Control*, *132*(November 2023), 104058. https://doi.org/10.1016/j.ijggc.2024.104058

Li, D., Peng, S., Guo, Y., Lu, Y., & Cui, X. (2021). CO2 storage monitoring based on time-lapse seismic data via deep learning. *International Journal of Greenhouse Gas Control*, *108*(November 2020), 103336. https://doi.org/10.1016/j.ijggc.2021.103336

Lu, J., Kharaka, Y. K., Thordsen, J. J., Horita, J., Karamalidis, A., Griffith, C., Hakala, J. A., Ambats, G., Cole, D. R., Phelps, T. J., Manning, M. A., Cook, P. J., & Hovorka, S. D. (2012). CO 2-rock-brine interactions in Lower Tuscaloosa Formation at Cranfield CO 2 sequestration site, Mississippi, U.S.A. *Chemical Geology*, *291*, 269–277. https://doi.org/10.1016/j.chemgeo.2011.10.020

Ringrose, P. (2023). *Storage of Carbon Dioxide in Saline Aquifers: Building confidence by forecasting and monitoring*. https://doi.org/10.1190/1.9781560803959

Xue, Y., Clarke, D., & Wang, K. (2023). Multi-scenario deep learning 4D Inversion: A Brazil pre-salt case study. *First Break*, *41*(9), 29–36. https://doi.org/10.3997/1365-2397.fb2023069

Zhang, R., Song, X., Fomel, S., Sen, M. K., & Srinivasan, S. (2013a). Case History: Time-lapse seismic data registration and inversion for CO2 sequestration study at Cranfield. *Geophysics*, *78*(6).

Zhang, R., Song, X., Fomel, S., Sen, M. K., & Srinivasan, S. (2013b). Time-lapse seismic registration and inversion for CO 2 sequestration study at Cranfield Part II : Pre- stack Analysis. *Annual Meeting SEG Houston 2013 Annual Meeting*, 5015–5020.