From Ideas to Deployment - A Joint Industry-University Research Effort on Tackling Carbon Storage Challenges with AI

Junjie Xu[1], Jiesi Lei[2], Yang Li[3], Junfan Ren[4], Jian Qiu[5], Biao Luo[5], Lei Xiao[5*], Wenwen Zhou[5*]

[1]Department of Energy and Power Engineering, Tsinghua University, Beijing, China
[2]School of Environment, Tsinghua University, Beijing, China
[3]Department of Electronic Engineering, Tsinghua University, Beijing, China
[4]College of Petroleum Engineering, China University of Petroleum (Beijing), Beijing, China
[5]Product and Solution & Website Business Unit, Alibaba Cloud, Hangzhou, Zhejiang, China
{xjj20, leijs20, yangli21}@mails.tsinghua.edu.cn, 19801212354@163.com, {qiujian.qj, luobiao.luo, xiaolei.xiao, zhoubo.zww}@alibaba-inc.com
[*]Corresponding author

Abstract

Carbon capture and storage (CCS) offers a promising means for significant reductions in greenhouse gas emissions and climate change mitigation at a large scale. Modeling CO_2 transport and pressure buildup is central to understanding the responses of geosystems after CO_2 injection and assessing the suitability and safety of CO_2 storage. However, numerical simulations of geological CO_2 storage often suffer from its multi-physics nature and complex non-linear governing equations, and is further complicated by flexible injection designs including changes in injection rates, resulting in formidable computational costs. New ideas have emerged such as data-driven models to tackle such challenges but very few have been fully developed and deployed as reliable tools. With the joint efforts of industry and universities, we are currently working on a new mechanism of fostering cross-disciplinary collaboration, developing, deploying, and scaling data-driven tools for CCS. A deep learning suite that can act as an alternative to CCS variable rate injection simulation will be the first tool developed under this mechanism. Based on the surrogate model, optimal design of injection strategy under pressure buildup constraints will be enabled with machine learning.

Introduction

Geological storage of CO_2 in saline aquifers, depleted oil and gas fields or unmineable coal seams, represents one of the most important processes for reducing anthropogenic emissions of greenhouse gases. The IEA's Sustainable Development Scenario proposes that over 90% of CO_2 captured from various sectors and sources should be destined for geological storage through to 2070 (IEA, 2020). CO_2 storage in deep saline aquifers, in particular, is a vital option with an estimated capacity of ~ 10^3-10^4 Gt globally (IPCC, 2005).

Successful CO_2 storage projects are necessitated by accurate prediction of geo-sequestration processes of CO_2 with regard to specific reservoir conditions and injection configurations. Numerical simulation is the primary tool used for predicting the flow transport of CO_2 by solving conservation equations. In such practices, the prediction of CO_2 transport is governed by highly non-linear partial differential equations (PDEs) and requires fine spatial and temporal discretization for accurate depiction of flow processes. Thus, large-scale numerical simulation is computationally expensive, inefficient and even infeasible for some CCS storage optimization problems.

Machine learning has recently shown a growing potential for applications to CCS problems as a substitute of conventional numerical simulation, with comparable fidelity as numerical simulations in some simple tasks, but at a much faster speed. For example, a rough set-based machine learning (RSML) technique was used to test the storage integrity of geological reservoirs (Aviso et al. 2019). We could also apply support vector machine (SVM) and random forest (RF) algorithms to predict CO_2 trapping efficiencies in saline formations (Thanh and Lee 2021). In more complicated cases, typical outputs (for example, CO_2 saturation distributions and pressure responses) from numerical simulators can be directly modeled by machine learning and deep learning methods. By means of Bayesian learning and principle component analysis (PCA), the spatial distribution of CO_2 concentration and pressure at the top of the reservoir can be forecast accurately (Lu et al. 2022). Furthermore, CCSNet developed by Wen et al. (2021) used a temporal 3D CNN model architecture that can predict the dynamic changes of CO_2 storage process in deep saline aquifers for single well injection scenarios with constant injection rates. Later, a Fourier neural operator-based deep-learning model (U-FNO) was proposed to extend state-of-art CNN models to address the effect of reservoir anisotropy on CO_2 storage under simplified CO_2 injection configurations, i.e., single well and constant rates (Wen et al. 2022). Such advances greatly improved the efficiency of dealing with multiphase transport problems associated with CO_2 storage, representing several of the successful trials at intersection of CCS and machine learning.

In industrial-scale carbon sequestration operations, injection designs (e.g., injection rate, injection scenario, well configuration, perforation interval) play a central role in determining CO_2 storage capacity, risks and liabilities, and ultimately the environmental and economic benefits of a project (Al-Khdheeawi et al., 2018). For example, the excessive pressure buildup from defective injection design such as a large and persisting injection rate may result in CO_2 leakage and environmental contamination with increasing "failure" risks of caprock fracturing, leakage up abandoned wells, and induced seismicity (Buscheck et al., 2012). Therefore, dynamical CO_2 injection strategy characterized by variable injections rates is the norm, and pressure buildup as a limiting factor for safe trapping of CO_2 must be considered. However, to date, surrogate models have not been designed and trained for dynamic CO_2 injection scenarios.

We seek to develop a deep-learning based toolkit that achieves rapid and accurate prediction of CO_2 plume migration processes under dynamic CO_2 injection schemes, as well as optimization of CO_2 injection strategy with regard to CO_2 storage capacity and security. To this end, we have integrated theoretical and computing resources from cloud technology providers and universities, adopted deep learning and machine learning methods, and proposed a lightweight surrogate model construction method based on mathematical empirical formula. This work is a pioneering attempt of interdisciplinary and industry-university collaboration. It is expected to appeal to both academia and industry, and a broad audience with interests in how to bring new ideas to deployment on tackling climate change with artificial intelligence.

Research Plan

Our research plan will be divided into four phases: data generation, surrogate model construction, machine learning-based optimization model research, and model deployment (Figure 1).

Phase 1: Data Generation

As a data-driven study, we first need to obtain data which describe the storage process of CO_2 in saline aquifers for subsequent model training and validation. A specific carbon storage scenario consists of several categories of parameters, including geological conditions, CO_2 injection conditions, etc. We will use numerical simulation methods to generate geological evolution data for a range of specific saline carbon storage scenarios. Specifically, we will consider the following categories of variable parameters to ensure a diversity of carbon sequestration scenarios:

- Reservoir conditions: basic information about geological formations, including initial pressure, temperature, reservoir thickness, etc.;
- Geological model: spatial distribution of permeability distribution in the reservoir;
- Rock properties: rock properties that affect carbon dioxide migration in the formation, including capillary pressure curves, relative permeability curves, etc.;
- CO_2 injection design: perforation position, perforation thickness, variable CO_2 injection rates, etc.;

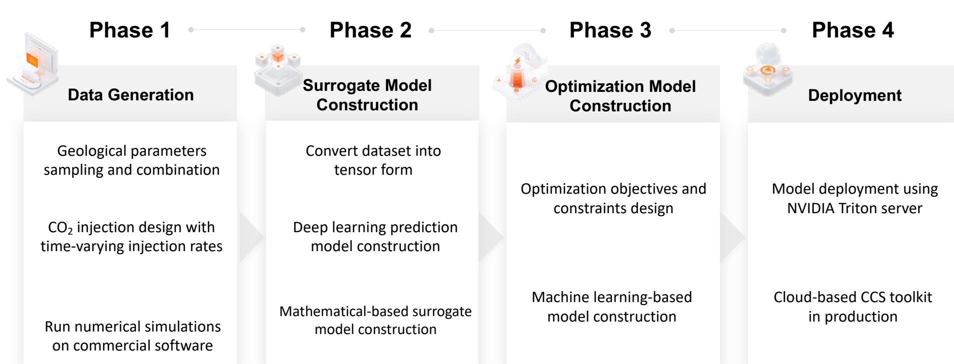

Figure 1: Research plan for surrogate and optimization model construction for CO_2 storage process in deep saline formations

We will sample the parameters above within a reasonable range and generate a series of diverse carbon storage scenarios in saline aquifers. It should be emphasized that the CO_2 injection rate under each scenario will vary over time, since our goal is to develop a surrogate prediction model for variable CO_2 injection rate conditions. At the same time, the setting of time steps will be non-uniform in order to take both transient and steady-state responses of geological evolution into account.

We will use a widely adopted numerical simulator to generate data of CO_2 migration and geological evolution within 30 years for each scenario. We plan to generate data of more than 10,000 scenarios through numerical simulation, and use these data as the training dataset for this study. As of the submission of this paper, Phase 1 is currently underway.

Phase 2: Surrogate Model Construction

Based on the data generated in the first phase, we will use data-driven methods to build surrogate models, which enable rapid and accurate prediction of CO_2 saline aquifer storage processes under specific scenarios.

Deep learning is the main approach we will take. We will convert the parameter inputs such as the initial pressure field and permeability map into matrices that can be recognized and understood by the deep learning framework. The results of the storage evolution processes can also be converted to matrix form. After that, a convolutional neural network-based prediction model can be constructed. Using our dataset, we will explore the performance of different deep learning models for dynamic CO_2 injection rate scenarios. In this process, some existing deep learning models with similar scope to this paper will also provide important inspiration for model construction.

In the process of building a deep learning model, based on the idea of multidisciplinary integration, the combination of network structure building and geological mechanism interpretation will also be the focus of our study. We will consider the thermodynamic and geological models involved in the carbon storage scenarios, and explore the possibility of utilizing prior knowledge. Great efforts will be made to improve the accuracy and interpretability of the model.

While deep learning models, if trained appropriately, are capable of predictions of a collection of variables related to CO_2 transport and distribution, physics-based surrogate models can be constructed for calculations of specific variables of interest at lower model training costs. For example, we have demonstrated the feasibility of applying the Duhamel's principle (Duhamel, 1833) from the well testing of oil production to solve the bottom hole pressure response of variable CO_2 injection rates with constant rate data (Figure 2). According to Duhamel's principle, when we know the constant-rate pressure responses, the bottom hole pressure response is given in discrete form as:

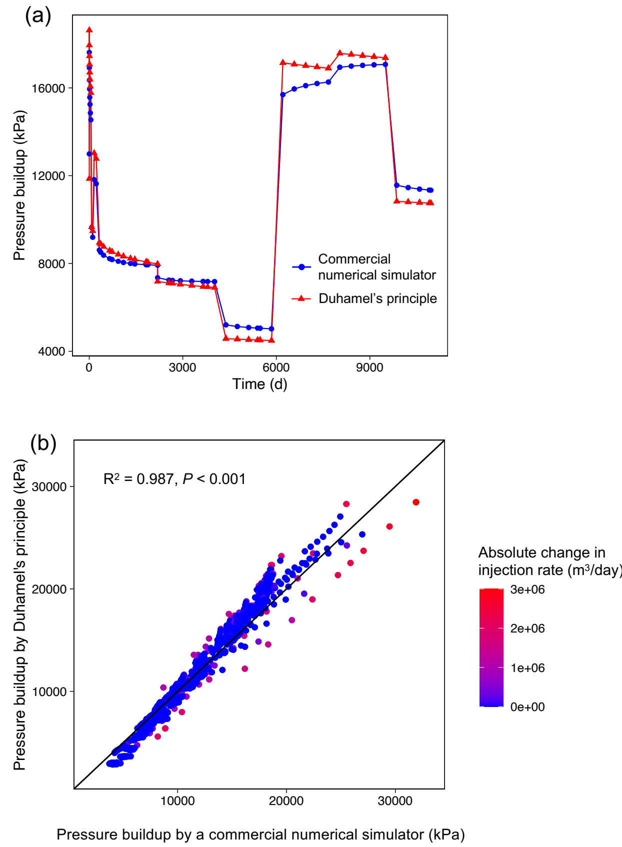

Figure 2: Application of the Duhamel's principle to solve the bottom hole pressure response of variable CO_2 injection rates with constant rate data. (a) A single case demonstration; (b) Parity plot showing the predictive capability of the physics-based method versus commercial simulator.

$$\Delta p(t) = \sum_{i=1}^{n} (q_i - q_{i-1})\big(p_u(t - t_{i-1})\big) \qquad (1)$$

where $\Delta p(t)$ is the pressure response at time t, q_i is the i-th injection rate in variable CO_2 injection rate series, and $p_u(t)$ is the constant-rate pressure response per unit injection rate at time t.

We tested the accuracy of calculations using Duhamel's principle under a series of carbon sequestration scenarios with different parameters. A total of 20 CO_2 injection scenarios were randomly generated with 10 injection rate changes over a 30-yr period. Predictions of time-series bottom hole pressure buildup based on Duhamel's principle using constant rate data aligned well with those by numerical simulations of variable injection schemes (Figure 2a). Parity plot showing the predictions based on Duhamel's principle versus the corresponding simulated results for each case at each time point provides a more complete picture of the predictive capabilities of the physics-based method (Figure 2b).

Duhamel's principle predicts the bottom hole pressure response exceedingly well ($R^2 = 0.987$, $P < 0.001$), suggesting that simple and lightweight surrogate models, though weaker in applicability compared with deep learning models, are also worth considering in the face of some specific problems.

Phase 3: Optimization Model Construction

The third phase will be to build an optimization model based on machine learning to find the optimal strategy for CO_2 injection under specific objectives and constraints. Currently, we hope to solve the optimal time series for CO_2 injection with the objective of maximizing the total amount of carbon storage within 30 years. The pressure at the injection well perforation and some other key locations will be constrained. The optimization objectives and constraints above can also be further adjusted according to other practical considerations. The surrogate model obtained in the second phase is expected to greatly improve the prediction efficiency and lay an indispensable foundation for the optimization work in this phase.

Since the CO_2 injection rate is a continuous variable, and its time series has infinite choices, we will discretize the value of injection rate and time. Subsequently, we will use machine learning methods to solve the optimization problem which might not be solved otherwise due to the large decision space. In this model, different CO_2 injection rates will be regarded as different behaviors of the "agent". We will formulate appropriate reward functions based on the total mass of injected CO_2 and local pressure buildup to help the model find the optimal action strategy in different scenarios.

In this process, we will explore machine learning methods such as reinforcement learning, genetic algorithm, particle swarm optimization, deep learning and other methods to improve the performance of the agent. As a consequence, we hope that the agent can efficiently and accurately find the optimal strategy under different carbon storage scenarios, that is, the optimal CO_2 storage solution that satisfies various constraints.

Through the above three phases of work, we will finally get an efficient solution to the problem of carbon storage evolution prediction and optimization, driven by artificial intelligence methods. Then we will be able to integrate all the results into a suite. The suite is expected to rapidly predict the geological evolution and CO_2 migration process under specific CO_2 saline aquifer storage scenario, within a few seconds at most. Furthermore, it can find the optimal solution of CO_2 injection rate time series for specific optimization objectives and constraints quickly, ensuring both efficiency and safety of the storage process.

Phase 4: Model Deployment

The final phase will be model deployment with cloud technology. We will deploy our surrogate model and optimization model using NVIDIA Triton inference server (https://developer.nvidia.com/nvidia-triton-inference-server). Triton is an open-source inference-serving software for deploying AI in applications. It can satisfy diverse application requirements and be compatible with AI models implemented in different ways. Using RAPIDS (https://developer.nvidia.com/rapids) as the inference backend, Triton's model analyzer intelligently determines the optimal model configuration, and combines concurrent execution and dynamic batching capabilities to improve inference efficiency. The model we deploy will provide users with flexible, scalable, and easy-to-access CCS guidance services powered by cloud technologies. Users will be able to customize input variable combinations, including geological attributes and injection configurations. Model outputs such as pressure buildup and gas saturation maps at fine spatial-temporal resolution will be provided within seconds. In the optimization service, critical thresholds for safe CO_2 trapping (e.g., maximum pressure buildup and maximum injection rates) can be customized. The optimization model will enable the search for a "best" combination of CO_2 injection rates at different time points that ensures the largest amount of CO_2 storage over the project cycle, while taking into account safety requirements specified by users. As such, the deployed models will be able to largely facilitate site selection and injection design of real-world CCS projects.

Future work

In the short run, we will focus on expanding the capabilities of the dynamic CO_2 storage surrogate model and exploring the optimal design of injection schemes under diverse constraints in engineering applications. Specifically, injection design optimization will take into account new constraints, such as reservoir pressure associated with fault activation. Future work will also be to extend the surrogate model approach that integrates mathematical models with machine learning to more CCS processes that rely on traditional numerical simulations. In the long-run, efforts will be directed toward the development of the mechanism that empowers the designing and scaling of data-driven tools for CCS by cross-disciplinary, industry-university collaboration.

Acknowledgements

The authors would like to acknowledge support from Professor Guangzhi Liao and Professor Liang Xue from China University of Petroleum (Beijing). Professor Shujuan Wang from Tsinghua University has also provided constructive advices. The authors are very grateful to share ideas and discussions with CCSNet's author Gege Wen from Stanford University. Hongxiang Tong from Alibaba Cloud has also provided valuable suggestions of model deployment best practices.

References

Al-Khdheeawi, E. A.; Vialle, S.; Barifcani, A.; Sarmadivaleh, M.; and Iglauer, S. 2018. Impact of injection scenario on CO_2 leakage and CO_2 trapping capacity in homogeneous reservoirs. In *Offshore Technology Conference Asia*. OnePetro. doi.org/10.4043/28262-MS.

Aviso K. B.; Janairo J. I. B.; Promentilla M. A. B.; and Tan R. R. 2019. Prediction of CO_2 storage site integrity with rough set-based machine learning. *Clean Technologies and Environmental Policy* 21: 1655-1664. doi.org/10.1007/s10098-019-01732-x.

Buscheck, T. A.; Sun, Y.; Chen, M.; Hao, Y.; Wolery, T. J.; Bourcier, W. L.; Court, B.; Celia., M. A.; Friedmann, S. J.; and Aines, R. D. 2012. Active CO_2 reservoir management for carbon storage: Analysis of operational strategies to relieve pressure buildup and improve injectivity. *International Journal of Greenhouse Gas Control* 6: 230-245. doi.org/10.1016/j.ijggc.2011.11.007.

Duhamel, J. M. C. 1833. Mémóire sur la méthode générale relative an mouvement de la chaleur dansles corps solides plonges dans les milieaux dont la température vaie avec te temps. *J. Ecole Polyt. Paris* 14: 20.

IEA, P., 2020. CCUS in Clean Energy Transitions. Technical Report. https://www. iea.org/reports/ccus-in-clean-energy-transitions.

IPCC. 2005. *Carbon Capture and Storage*, Bert Metz, Ogunlade-Davidson, Heleen de Coninck, Manuela Loos and Leo Meyer (Eds.) Cambridge University Press, UK. pp 431. Available from Cambridge University Press, The Edinburgh Building Shaftesbury Road, Cambridge CB2 2RU ENGLAND.

Lu D.; Painter S. L.; Azzolina N. A.; Burton-Kelly M.; Jiang T.; and Williamson C. 2022. Accurate and rapid forecasts for geologic carbon storage via learning-based inversion-free prediction. *Frontiers in Energy Research* 9: 752185. doi.org/10.3389/fenrg.2021.752185.

Thanh H. V. , Lee K. K. 2021. Application of machine learning to predict CO_2 trapping performance in deep saline aquifers. *Energy* 239: 122457. doi.org/10.1016/j. energy.2021.122457.

Wen, G.; Hay, C.; and Benson, S. M. 2021. CCSNet: a deep learning modeling suite for CO_2 storage. *Advances in Water Resources* 155: 104009.

Wen, G.; Li, Z.; Azizzadenesheli, K.; Anandkumar, A.; and Benson, S. M. 2022. U-FNO—An enhanced Fourier neural operator-based deep-learning model for multiphase flow. *Advances in Water Resources* 163: 104180. doi.org/10.1016/j.advwatres.2022.104180.