Convolutional-Recurrent Proxy Model for Spatiotemporal CO_2 Monitoring

Misael M. Morales[1*], Carlos Torres-Verdín[1,2] and Michael J. Pyrcz[1,2]

[1]Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX, USA.
[2]Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA.

*Corresponding author(s). E-mail(s): misaelmorales@utexas.edu;

Abstract

The abstract serves both as a general introduction to the topic and as a brief, non-technical summary of the main results and their implications. Authors are advised to check the author instructions for the journal they are submitting to for word limits and if structural elements like subheadings, citations, or equations are permitted.

Keywords: Spatiotemporal forecasting, Convolutional neural network, Recurrent neural network, Proxy model

1 Introduction

Geologic CO2 sequestration (GCS) has emerged as a proven technology to reduce anthropogenic greenhouse gas emissions to the atmosphere [citation]. This has become increasingly popular worldwide due to the need to meet international climate protection agreements [citation]. However, there are several technical challenges associated with the modeling of large-scale GCS operations. In order to accurately forecast and monitor subsurface multiphase flow, physics-based high-fidelity numerical simulation is required. These numerical simulations are computationally intensive and time-consuming since they require iterative solutions of large-scale nonlinear systems of equations [citation]. Similarly, due to the large degree of uncertainty in subsurface data collection, inherent uncertainty in the spatial distribution of the properties of heterogeneous porous media require a

robust probabilistic assessment for improved engineering decision-making [citation]. In order to capture the fine-scale multiphase flow behavior given an uncertain spatial distribution of subsurface properties, a large number of forward numerical simulation runs are required, leading to very high computational costs [citation]. To overcome this, machine learning techniques have emerged as candidate reduced-order models (ROMs) for efficient parameterization and prediction of subsurface flow and transport behavior [citation].

Recent advancements in computing power, specifically GPU-enabled neural network models, have accelerated the fields of forward and inverse modeling [citation]. Classical techniques are often hindered by the size of the models and data, specifically the volume, velocity, variety, value, and veracity encountered in big data [citation]. By analyzing extensive data sets, machine learning techniques can uncover complex latent

1

patterns and relationships that may not be discernible through traditional methods [citation]. When combined with a reduced-order modeling framework, machine learning approaches can efficiently and accurately exploit latent or salient features hidden in the data, removing redundancies or noise, and decreasing the order of the problem significantly [citation]. These approaches can often be divided into two main categories, namely purely data-driven mapping operators or physics-informed neural networks (PINNs). Typically, the training process for PINNs is done by the minimization of the (physical) loss from the residual of the governing partial differential equations (PDEs) that govern the system along with the losses associated with the initial and boundary conditions [citation]. However, over variants of PINNs such as physics-guided or physics-constrained neural networks have also proven useful for subsurface energy resource engineering applications [citation]. On the other hand, data-driven mapping operators, or proxy models, are neural network architectures trained with labeled data that produce a mapping from input features to output parameters [citation]. This procedure requires significant amounts of training data but can be applied to a wide variety of settings and conditions [citation] but suffer from lack of generalization and struggle to provide accurate predictions away from the domain of the training data. For both approaches, typically, spatial relationships are captured through convolutional neural networks (CNNs) and the temporal relationships through recurrent neural networks (RNNs) [citation], but recent advancements in transformer-based architectures are showing improved performance compared to the aforementioned techniques [citation]. In general, efficient compression of the input features into a representative latent space is proven as an effective approach for spatial and temporal parameterization of the forward or inverse problem.

A number of machine learning-based proxy (or surrogate) models have been developed to estimate the reservoir behavior in subsurface energy resource applications. Most techniques rely on the concept of image translation, or pix2pix, where a target image is predicted from an input image [citation]. Maldonado and Pyrcz [1] developed a convolutional U-net model to predict pressure and saturation states given an uncertain geologic realization. This work is an example of image-to-image static forecasting, where the time state is given as an input, and the proxy model will predict a single response state of pressure and saturation at the given time. Wen and Benson [citation] developed a Fourier Neural Operator (FNO) architecture to predict image-to-image response states of pressure and saturation from an uncertain geologic realization and was further extended for multi-scale and nested domains [citation]. Moreover, numerous other proxy models have been developed for subsurface applications using more complex architectures such as generative adversarial networks (GANs) [citation] and transformers [citation]. However, most of these formulations are presented as an even-determined or sometimes over-determined estimation problem, with equal or greater number of input features compared to the output parameters since they are based on the pix2pix, or image-to-image formulation.

Moving beyond image-to-image predictions, Kim and Durlofsky [citation] developed a convolutional-recurrent proxy for image-to-series forecasting and discussed its advantages for closed-loop reservoir management under geologic uncertainty. This method moves beyond the image-to-image forecasting and exploits a spatiotemporal latent space in the encoder-decoder neural network architecture to obtain well flow rates and pressures over time from a static geologic realization. The image-to-series formulation can still be an even- or over-determined estimation problem, where we have equal or more inputs than outputs. Furthermore, Tang et al. [citation] and Jiang and Durlofsky [citation] developed a recurrent residual U-net (R-U-net) proxy for the prediction of dynamic pressure- and saturation-over-time from uncertain geologic realizations. This method aim to obtain dynamic response states over time from a single static input. This proxy is formulated as a more interesting under-determined estimation problem, where the number of input features is a fraction of the number of output parameters. However, the recurrent R-U-net proxy is limited by the fact that only the latent space receives spatiotemporal processing, while the model reconstruction is done via time-distributed deconvolutions, treating time as an additional "spatial" dimension, and not fully exploiting the spatiotemporal relations in the data

and latent space as an image-to-video forecasting formulation.

The problem of image-to-video forecasting has been approached previously by researchers in the field of computer vision. Iliadis et al. [citation] were the first to develop a deep learning-based framework for video compressive sensing to reconstruct a video sequence from a single measured frame using a deep fully-connected neural network, or artificial neural network (ANN). Despite excellent accuracy in the video predictions, this method is still limited by time-distributed fully-connected layers in the encoder and decoder portions of the network, thus not exploiting the spatiotemporal relationships in the data. Xu and Ren [citation] developed a three-part encoder-recurrent-decoder network for video reconstruction from the estimated motion fields of the encoded frames. The implementation is similar to that of Tang et al. [citation] and Jiang and Durlofsky [citation] in that it applies a recurrent update in the latent space but relies on time-distributed deconvolutions for the video frames reconstruction. Dorkenwald et al. [citation] developed a conditional invertible neural network (cINN) as a bijective mapping between image and video domains using a dynamic latent representation. The cINN architecture allowed for video-to-image and image-to-video predictions, proving possible the generation of video frames from a static input image. Finally, Holynski et al. [citation] implemented the idea of Eulerian motion fields to define the moving portions of an image and thus were the able to accurately reconstruct a series of video frames from a static image using a spatiotemporal latent space parameterization. These advancements in the field of computer vision and video compressed sensing serve as a foundation for our image-to-video spatiotemporal proxy model.

In this work, we propose a novel image-to-video spatiotemporal proxy model for the prediction of dynamic reservoir behavior over time from an uncertain static geologic realization. In this work, we apply the spatiotemporal proxy to a large-scale GCS operation. Our model exploits the spatial and temporal structures in latent space to dynamically reconstruct the time-dependent pressure and saturation states from a static geologic realization. The encoder portion of the network receives as inputs the static geologic

realization with channels representing the porosity, permeability, and facies distributions, and the location of CO2 injection well(s). The uncertain geologic realizations are generated from a wide array of possible geologic scenarios (e.g., fluvial, turbidite, and deep water lobe systems), and the number and location of CO2 injection wells is also considered uncertain. The model then reconstructs the dynamic pressure and saturation distributions using a spatiotemporal decoder network with convolutional long short-term memory (convLSTM) layers, which are concatenated with the residuals of the spatial latent parameterizations from the encoder network. Thus, it is not an encoder-recurrent-decoder architecture, but instead a fully spatiotemporal convolutional-recurrent image-to-video model. Our proxy model shows significant advantages compared to image-to-image and encoder-recurrent-decoder models in terms of computational efficiency and prediction accuracy and can be used as a replacement for high-fidelity simulations (HFS) in GCS projects as an image-to-video mapping operator.

In the methodology section, we discuss the proposed spatiotemporal proxy model architecture as well as the geologic modeling and numerical reservoir simulation steps required to generate the training data. In the results and discussion sections, we evaluate the training and performance of the proposed proxy model and compare its efficiency and accuracy to high-fidelity numerical simulations using a 2D synthetic case for large-scale GCS operations.

2 Methodology

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

2.1 Reservoir Model Description

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

2.2 Reservoir Simulation

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

Sample body text. Sample body text. Sample body text.

2.3 Proxy Model Architecture

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

2.3.1 Spatial Structure

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

2.3.2 Temporal Structure

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

3 Results

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

4 Discussion

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.
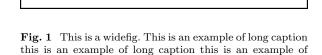The input format for the above table is as follows:

Table 1 Caption text

Column 1	Column 2	Column 3	Column 4
row 1	data 1	data 2	data 3
row 2	data 4	data 5^1	data 6
row 3	data 7	data 8	data 9^2

Source: This is an example of table footnote. This is an example of table footnote.

[1]Example for a first table footnote. This is an example of table footnote.

[2]Example for a second table footnote. This is an example of table footnote.

Fig. 1 This is a widefig. This is an example of long caption this is an example of long caption this is an example of long caption this is an example of long caption

5 Conclusions

In this study, we developed a deep learning-based spatiotemporal proxy model to provide flow predictions for a large-scale GCS operation. The key extension introduced is the use of a spatiotemporal convolutional-recurrent architecture for dynamic predictions of CO2 pressure and saturation distributions over time from an uncertain static geologic realization. The framework was developed as an image-to-video prediction, which is a noteworthy under-determined estimation problem. Specifically, the implementation extends the architectures of current encoder-recurrent-decoder models and provides a fast and accurate proxy as a replacement for high-fidelity numerical reservoir simulation.

The encoder block is composed of separable convolutions, squeeze and excite layers, and instance normalization. These three special implementations allows for precise parameterization of the geologic realization into a latent representation, without mixing the effects of Gaussian distributed properties against binary of binomial distributed properties. Using recursive convLSTM layers, the recurrent decoder block recursively predicts each dynamic state, or frame, from the concatenation of the previous latent representation and the intermediate encoding parameterizations. Thus, this architecture presents the proxy

as an image-to-video prediction formulation for dynamic reservoir states from a static geologic realization.

The spatiotemporal proxy was applied to a synthetic 2D GCS project with multiple uncertain geologic scenarios and random number and location of injection well(s). A total of 1,000 geologic models were obtained from a variety of possible geologic scenarios including fluvial, turbidite, and deep water lobe systems. The spatial distribution of porosity, permeability and facies, and the spatial location of the injector well(s) were used as the input data. The proxy then predicts the dynamic reservoir response over time, namely the video frames, corresponding to the dynamic CO_2 pressure and saturation distributions, which are obtained offline for training using HFS. The total training time is 78 minutes on a single NVIDIA Quadro M6000 GPU, and predictions are obtained with 98% accuracy within approximately 5.4 milliseconds, compared to the approximate 30 seconds required for HFS – a 6,000x speedup.

There are several possible directions that could be considered for future work. Although the spatiotemporal convolutional-recurrent proxy was only trained for CO_2 sequestration, it should be applicable for a range of processes such as compositional, geothermal, or conventional oil and gas systems. The proxy could also be applied to several subsurface energy resource workflows such as optimization and history matching. Moreover, it would be interesting to extend the proxy from a data-driven mapping operator to a PINN by including the discretized form of the governing PDE in the loss function and minimizing the residuals. Furthermore, the proxy is robust to uncertain geology and variable number and placement of injector wells but could be extended to variable well controls and applied to robust optimization and closed-loop reservoir management workflows.

Declarations. The authors declare no conflict of interests.

References

[1] Maldonado-Cruz, E., Pyrcz, M.J., 2022. Fast evaluation of pressure and saturation predictions with a deep learning surrogate flow model. Journal of Petroleum Science and Engineering 212, 110244. doi:10.1016/J.PETROL.2022.110244.