**Convolutional-Recurrent Proxy Model for Spatiotemporal CO2 Monitoring**

Misael M. Morales1\*, Carlos Torres-Verdín1,2, and Michael J. Pyrcz1,2

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

\*Corresponding author(s). E-mail(s): [misaelmorales@utexas.edu](mailto:misaelmorales@utexas.edu)

**Abstract**

*Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.*

**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 for efficient parameterization and prediction of subsurface flow and transport behavior [citation].

Recent advancements in computing power and GPU-enabled neural network models have accelerated the field 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 patterns and relationships that may not be discernible through traditional methods [citation]. When combined with reduced-order modeling (ROM), 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 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]. Typically, spatial relationships are captured through convolutional neural networks (CNNs) and the temporal relationships through recurrent neural networks (RNNs) [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.

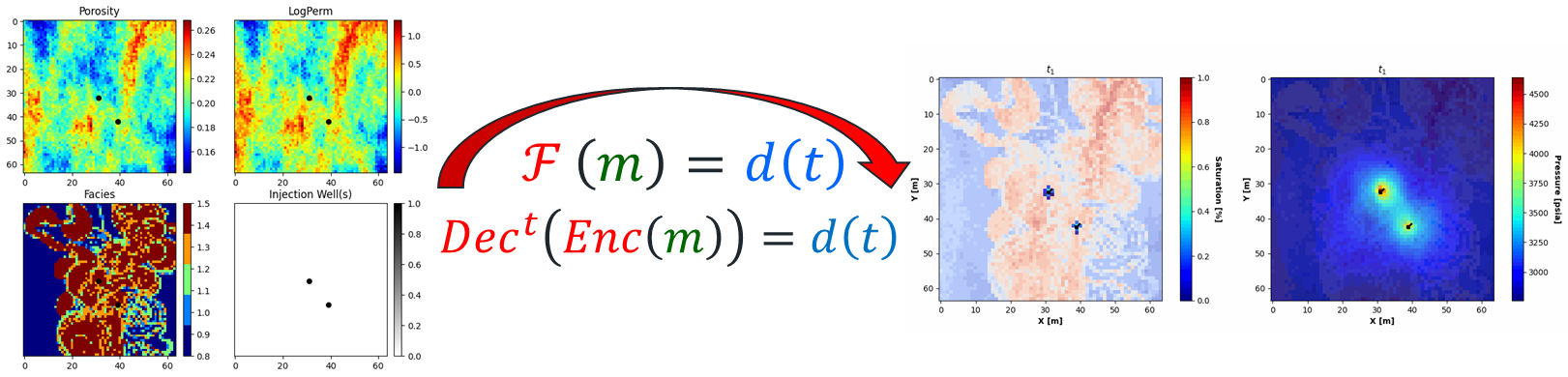
A number of machine learning-based proxy (or surrogate) models have been developed to estimate the reservoir behavior in subsurface energy resource applications. Most techniques rely on the concept of image translation, or pix-2-pix, where a target image is predicted from an input image [citation]. Maldonado and Pyrcz [citation] 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 dynamic state of pressure and saturation at the given time. Similarly, Wen and Benson [citation], developed a Fourier Neural Operator (FNO) architecture to predict image-to-image states of pressure and saturation from an uncertain geologic realization, and further extended the FNO proxy for multi-scale and nested domains [citation]. 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 exploit a spatiotemporal latent space in the encoder-decoder neural network architecture to obtain well flow rates and pressure over time from a static geologic realization. Moving beyond image-to-image and image-to-series, Tang et al. [citation] and Jiang and Durlofsky [citation] developed a recurrent residual U-net (R-R-U-net) proxy for the prediction of dynamic pressure- and saturation-over-time from uncertain geologic realizations. This was the first application of image-to-video predictions for subsurface flow modeling. However, the R-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 convolutions, treating time as an extra spatial dimension.

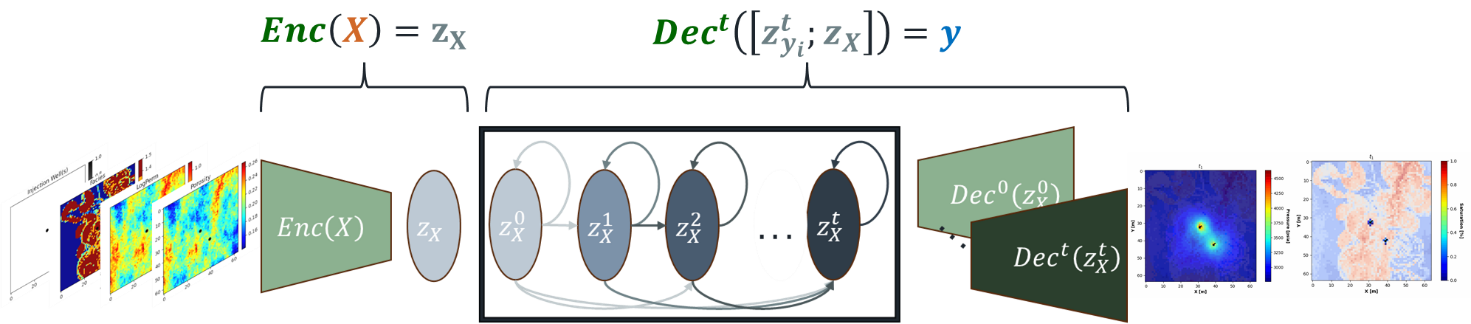
The problem of image-to-video forecasting has been approached previously by researchers in the field of computer vision. Iliadis et al. [citation] were the first to develop a deep fully-connected neural network for video compressive sensing to reconstruct a video sequence from a single measured frame. Xu and Ren [citation] developed a three-part encoder-recurrent-decoder network for video reconstruction from the estimated motion fields of the encoded frames. Holynski et al. [citation] implemented the concept of Eulerian motion fields to define the moving portions of an image and reconstructed a series of video frames. Finally, Dorkenwald et al. [citation] developed a conditional invertible neural network as a bijective mapping between image and video domains using a dynamic latent representation. These advancements in the field of computer vision and video compressed sensing serve as a foundation for our image-to-video spatiotemporal proxy model.

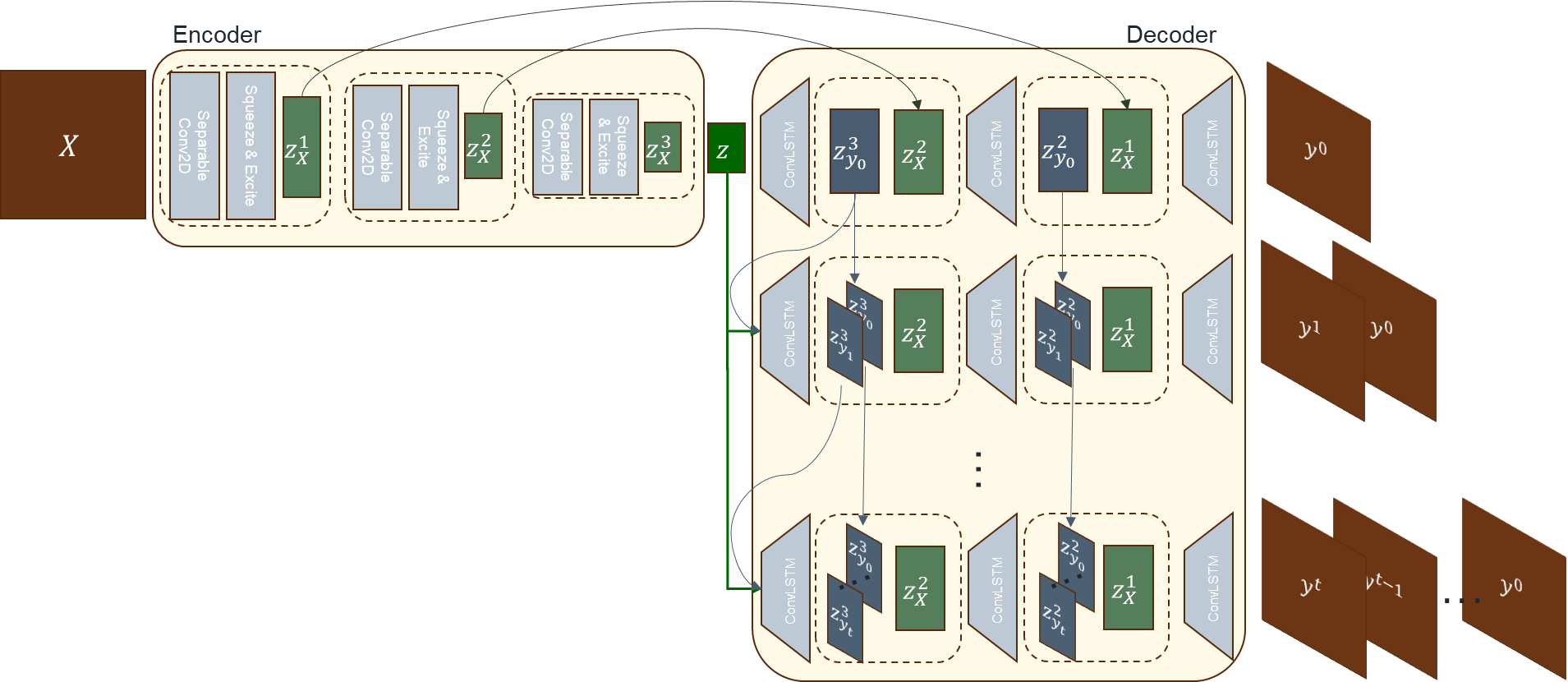
In this work, we propose a novel image-to-video spatiotemporal proxy model for the prediction of dynamic reservoir behavior over time from an uncertain static geologic realization. Our model exploits the spatial and temporal structures in latent space to dynamically reconstruct the time-dependent pressure and saturation states from a static geologic realization. In this work, we apply the spatiotemporal proxy to a large-scale GCS operation. The proxy receives as input a static geologic realization with four main inputs, the porosity, permeability, and facies distributions, and the location of CO2 injector well(s). The model then reconstructs the dynamic pressure and saturation plume migration over time. The uncertain geologic realizations are generated from a wide array of possible geologic scenarios (e.g., fluvial, turbidite, and deepwater lobe systems), and the number and location of CO2 injection wells is also considered uncertain. Our proxy model shows significant advantages to image-to-image and previous image-to-video models in terms of computational efficiency and prediction accuracy and can be used as a replacement for high-fidelity simulations in GCS projects.

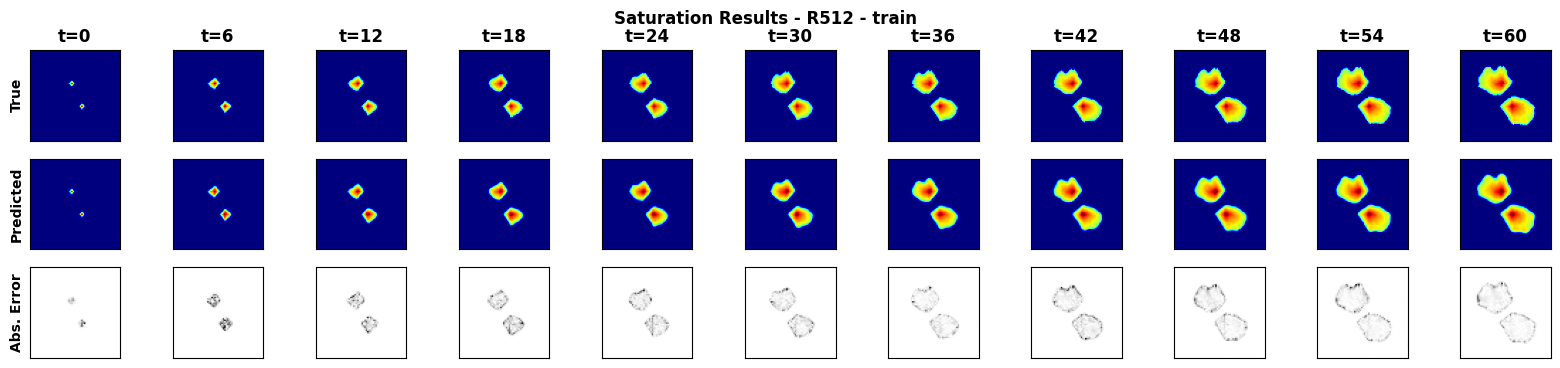
In the methodology section, we discuss the proposed spatiotemporal proxy model architecture as well as the geologic modeling and numerical reservoir simulation steps required to generate the training data. In the results and discussion sections, we evaluate the performance of the proposed proxy model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

1. **Methodology**
   1. Reservoir model description
   2. Reservoir simulation
   3. Model architecture
      1. Spatial structure
      2. Temporal structure
2. **Results and Discussion**
   1. Training performance
   2. Prediction results
   3. Cumulative co2 as a proxy for PVI
3. **Conclusions**









A screenshot of a computer generated image

Description automatically generated