

Contents lists available at ScienceDirect

International Journal of Greenhouse Gas Control



journal homepage: www.elsevier.com/locate/ijggc

Deep-learning-based coupled flow-geomechanics surrogate model for CO₂ sequestration

Meng Tang^{*, a}, Xin Ju^b, Louis J. Durlofsky^a

^a Department of Energy Resources Engineering, Stanford University, Stanford, CA 94305, United States
^b Atmospheric, Earth, and Energy Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, United States

А	R	Т	I	С	L	Е	Ι	Ν	F	0
---	---	---	---	---	---	---	---	---	---	---

Keywords: Carbon storage Coupled flow and geomechanics Surrogate model Deep learning Data assimilation History matching

ABSTRACT

A deep-learning-based surrogate model capable of predicting flow and geomechanical responses in CO₂ storage operations is presented and applied. The 3D recurrent R-U-Net model combines deep convolutional and recurrent neural networks to capture the spatial distribution and temporal evolution of 3D saturation and pressure fields and 2D surface displacement maps. The method is trained using high-fidelity simulation results for 2000 storage-aquifer realizations characterized by multi-Gaussian porosity and log-permeability fields. Detailed comparisons between surrogate model and full-order simulation results for new storage-aquifer realizations are presented. The saturation, pressure and surface displacement fields provided by the surrogate model display a high degree of accuracy, for both individual realizations and ensemble statistics. The recurrent R-U-Net surrogate model is applied with a rejection sampling procedure for data assimilation. Although the (synthetic) observations consist of only a small number of surface displacement measurements, significant uncertainty reduction in pressure buildup at the caprock is achieved.

1. Introduction

Carbon capture, utilization and storage (CCUS) is expected to play an important role in reducing greenhouse gas emissions to the atmosphere. In CCUS, supercritical CO₂ is injected into deep brine aquifers or abandoned oil reservoirs for permanent storage. According to a recent assessment by the International Energy Agency (International Energy Agency, 2020), in the Sustainable Development Scenario, a total of 650 megatonnes (Mt) per year of anthropogenic CO₂ is required to be geologically sequestered by 2030. This increases to 5266 Mt/year by 2050. Current large-scale commercial CCUS projects have a capacity of only about 40 Mt/year. Thus, worldwide we will need about 150 projects, at the 4 Mt/year scale (which is the scale considered in this work) by 2030 to satisfy the Sustainable Development Scenario. Fast and accurate predictions of flow and geomechanics will be required in order to effectively manage these large-scale storage operations.

The modeling of the geomechanical response of the formation is very important in CCUS operations because excessive pressure buildup can induce fracturing of the caprock, or activate pre-existing faults, through which CO_2 or brine can leak. The numerical simulation of coupled flow and geomechanics problems can be computationally prohibitive, however. This leads to significant challenges in the practical application of data assimilation (history matching) and uncertainty quantification, both of which require large numbers of forward simulations. Effective surrogate models for CO_2 storage will therefore be very useful in these settings.

Our goal in this work is to develop a deep-learning-based surrogate model to treat coupled flow and geomechanics in carbon storage operations. The resulting model will then be applied for data assimilation, using surface displacement data, to estimate aquifer properties and thus reduce uncertainty in pressure buildup at the caprock. The 3D surrogate model, which is trained to predict pressure and CO_2 saturation throughout the storage aquifer along with surface (ground-level) vertical deformation, represents an extension of our recently developed 2D and 3D surrogate models for two-phase subsurface flow in the absence of geomechanical effects (Tang et al., 2020; 2021).

A number of investigators have developed and applied numerical models for coupled flow and geomechanics in carbon storage operations (Fu et al., 2021; Fuchs et al., 2019; Ju et al., 2021; Li and Laloui, 2016; Shi et al., 2013; Talebian et al., 2013; Vilarrasa et al., 2010). Vilarrasa et al. (2010) modeled an axisymmetric reservoir-caprock system using hydromechanical coupling to evaluate the probability of reactivation or

* Corresponding author. E-mail addresses: mengtang@stanford.edu (M. Tang), ju1@llnl.gov (X. Ju), lou@stanford.edu (L.J. Durlofsky).

https://doi.org/10.1016/j.ijggc.2022.103692

Received 8 August 2021; Received in revised form 10 February 2022; Accepted 3 May 2022 1750-5836/ $\$ 2022 Elsevier Ltd. All rights reserved.

the creation of fractures, while Eshiet and Sheng (2014) introduced a discrete element method for the coupled problem. Shi et al. (2013) conducted a simulation study to match InSAR surface uplift data for the In Salah (Algeria) CO_2 storage project. Li and Laloui (2016) also studied the In Salah project, using a model that included flow, geomechanical and thermal effects. Their goal was to predict surface deformation and assess the effects of Biot's coefficient and temperature. Additional studies involving coupled flow and geomechanics in CO_2 storage settings were reviewed by Rutqvist et al. (2019). In this work we apply a computational framework that couples multiphase flow and geomechanics to model poroelastic rock response during CO_2 injection. The numerical approach entails a sequential-implicit fixed-stress scheme to couple a multiphase flow module and a mechanics module (Fu et al., 2021; Kim et al., 2011).

Surrogate models, also referred to as proxy models, can be used to replace computationally expensive forward models in problems requiring many related simulation runs. Example application areas include data assimilation, optimization, and uncertainty quantification. Surrogate models can be classified into simplified physical/numerical models and statistical (e.g., deep learning) methods. Reduced-order numerical models based on proper orthogonal decomposition (POD), for example, have been developed for problems involving coupled flow and geomechanics by Florez and Gildin (2019) and Jin et al. (2020). Neither of these studies, however, was performed in the context of geological carbon storage.

Deep-learning-based procedures approximate the input-output relationship of simulation data using statistical tools. Tang et al. (2020, 2021) developed recurrent R-U-Net surrogate models, for 2D and 3D two-phase subsurface flow problems, which combine a residual U-Net with convLSTM networks to predict the evolution of the saturation and pressure fields. A residual U-Net was additionally used, in an autoregressive strategy, to predict flow response in 2D problems with varying well-control specifications (Jiang et al., 2021). Deep-learning-based surrogate models have also been implemented for CO2 storage problems. Mo et al. (2019) developed a deep neural network framework to forecast CO₂ plume movement for random 2D permeability fields, while Wen et al. (2021) devised a general framework for predicting CO₂ plume location in axisymmetric (single-well) scenarios. There have been a number of other deep neural network models for subsurface flow and transport in applications other than CO₂ storage. Some of these studies are discussed by Tang et al. (2020, 2021). To our knowledge, the development of a deep-learning-based surrogate for CO2 storage with coupled flow and geomechanics, which is a key goal of this work, has yet to be accomplished.

The prediction of the pressure field and plume location, and the identification of leaks, is important for the assessment of risk in CO2 storage operations. Uncertainty in these essential quantities can be reduced through use of surveillance/monitoring data in conjunction with data assimilation procedures. Several studies (Cameron et al., 2016; González-Nicolás et al., 2015; Jung et al., 2015), for example, have used pressure monitoring data for leak detection. Chen et al. (2020) used pressure and CO₂ saturation data at monitoring wells to reduce uncertainty in plume location. Seismic data can also be incorporated for storage-aquifer characterization (Liu and Grana, 2020). The models used in these papers did not include coupled flow and geomechanics. There have, however, also been data assimilation studies for CO₂ storage that included geomechanical effects. de la Torre Guzman et al. (2014) considered compartmentalized aquifers and estimated fault transmissibility and near-well permeability using injection well pressure and surface deformation data. Jahandideh et al. (2021) used microseismic data to estimate rock properties including permeability and Young's modulus. Both of these studies took properties to be constant within a region/compartment, and both applied ensemble-based methods for the inversion.

A number of researchers have addressed data assimilation in other types of coupled flow-geomechanics systems. Wilschut et al. (2011) used surface subsidence and production data, with an ensemble Kalman filter method, to estimate fault transmissibilities in a pseudo-3D gas-bearing compartmentalized reservoir. Similarly, Jha et al. (2015) applied an ensemble smoother (ES) to estimate rock properties in a 3D homogeneous gas reservoir by assimilating both surface deformation and well data. ES was also applied by Zoccarato et al. (2016) to estimate geomechanical parameters such as Poisson's ratio in a 3D homogeneous gas storage model using surface displacement data. Tang (2018) applied both derivative-free optimization and an ES procedure to invert for (layered) permeability and Young's modulus in 2D models using well production data and surface displacement data. These studies, however, only involved the estimation of a limited number of parameters in homogeneous or layered models; i.e., detailed rock property fields were not considered. In addition, the data assimilation algorithms used in the studies with 3D models were restricted to ensemble-based methods because of the high computational demands associated with the forward modeling. In recent work, Alghamdi et al. (2021, 2020) developed an adjoint-gradient inversion procedure to determine permeability fields in groundwater-aquifer settings from surface deformation data. The permeability fields considered in these studies were areally heterogeneous (and characterized by Gaussian fields), but were assumed to have no variation in the vertical direction.

In this study, we extend the 3D recurrent R-U-Net surrogate model developed by Tang et al. (2021) to model coupled flow and geomechanics in CO₂ storage settings. The problem domain in the high-fidelity solution includes the storage aquifer, a surrounding region, overburden (which extends up to the Earth's surface), and bedrock. Extensive domains of this type are required for the proper modeling of geomechanical effects. The storage aquifer is characterized by random 3D multi-Gaussian permeability and porosity fields of prescribed correlation structure. The surrogate model is trained to predict the 3D saturation and pressure fields in the storage aquifer as well as the ground (surface-level) 2D vertical displacement field. We apply the overall framework to solve an inverse problem in which synthetic surface deformation data are used to predict the detailed permeability and porosity fields in the aquifer. A rigorous rejection sampling algorithm is applied. The resulting posterior models are then used to predict pressure buildup at the top of the storage reservoir.

This paper proceeds as follows. In Section 2, we first present the coupled flow and geomechanics equations applicable for CO_2 sequestration. We then describe the recurrent R-U-Net surrogate model and associated data-processing and training procedures. In Section 3, the 3D surrogate model is applied for coupled flow (CO_2 -brine) and geomechanics problems. In the simulations, CO_2 injection is accomplished via four vertical wells. Our specific interest is in predicting solutions for new geomodels characterized by multi-Gaussian random fields. A detailed assessment of surrogate model accuracy is presented. Next, in Section 4, the trained surrogate model is applied for data assimilation. We conclude in Section 5 with a summary and suggestions for future research directions.

2. Governing equations and surrogate model

In this section, we present the governing equations for coupled multiphase flow and geomechanics used to model CO_2 storage operations. The general simulation setup and solution variables are then discussed. Next, the recurrent R-U-Net surrogate model for the coupled system, along with the training process and data preprocessing, are described.

2.1. Governing equations for coupled problem

In the formulation for the coupled problem, the multiphase fluid and solid skeletons are treated as overlapping continua. The governing equations for the flow problem involve statements of mass conservation for water and CO_2 components, which flow in two phases, referred to as

the aqueous and gaseous phases. The conservation equations are expressed as:

$$\nabla \cdot \left(\sum_{j} \rho_{j} x_{j}^{r} \mathbf{v}_{j}\right) + q^{r} = \frac{\partial}{\partial t} \left(\sum_{j} \phi \rho_{j} S_{j} x_{j}^{r}\right), \tag{1}$$

where subscript j = a indicates the aqueous phase and j = g the gaseous phase, superscript r = w denotes the water component and r = c the CO₂ component, ρ_j is phase density, x_j^r is the mass fraction of component r in phase j, \mathbf{v}_j is the Darcy velocity (given below), q^r represents the source/sink term, t is time, ϕ is true porosity (defined as the ratio of the pore volume to the bulk volume in the deformed configuration), and S_j is the phase saturation.

The Darcy velocity for phase j, v_i , is given by

$$\mathbf{v}_{j} = -\frac{\mathbf{k}k_{rj}(S_{j})}{\mu_{j}} (\nabla p_{j} - \rho_{j}g\nabla z), \qquad (2)$$

where **k** is the absolute permeability tensor, k_{rj} , μ_j and p_j denote relative permeability, viscosity and fluid pressure of phase *j*, respectively, *g* is gravitational acceleration, and *z* is depth. The phase pressures are related to each other through the capillary pressure p_c ; i.e., $p_c = p_g - p_a$, with $p_c(S_g)$ a prescribed function. The system is closed by enforcing $S_g + S_a = 1$.

The governing quasi-static geomechanical equations are derived from the linear-momentum balance in the matrix, expressed as

$$\nabla \cdot \boldsymbol{\sigma} + \rho_m g \nabla z = 0. \tag{3}$$

Here σ denotes the Cauchy stress tensor and the composite matrix density ρ_m is a volume weighted average of phase densities ρ_j and rock density $\rho_{\rm rock}$, i.e.,

$$\rho_m = \phi \sum_j S_j \rho_j + (1 - \phi) \rho_{\text{rock}}.$$
(4)

From poroelasticity theory (Biot, 1941) and the assumption of small linearly elastic deformation, σ can be related to fluid pressure and solid displacement via:

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\epsilon} - bp_e \mathbf{1},\tag{5}$$

where **C** is a fourth-order stiffness tensor for the solid skeleton (associated with the drained-isothermal elastic moduli), ϵ denotes the strain tensor, with $\epsilon = \frac{1}{2}(\nabla \mathbf{d} + \nabla^{\mathsf{T}}\mathbf{d})$ and **d** the displacement vector, *b* is Biot's coefficient, p_e denotes the equivalent pore pressure derived from $p_e = \sum_i S_i p_i$ (Coussy, 2004), and **1** is the second-order identity tensor.

The fluid mass balance Eq. (1) and geomechanics momentum conservation Eq. (3) are tightly coupled through poromechanics (Biot, 1941; Coussy, 2004). This coupled flow and geomechanics system is solved using a fixed-stress sequential implicit procedure (Kim et al., 2011), which has been proven to yield accurate results with unconditional stability. In the sequential implicit scheme, the multiphase flow and geomechanical problems are solved iteratively, by updating the true porosity ϕ from its previous state through application of:

$$d\phi = \frac{b - \phi}{K_{dr}} (dp_e + d\sigma_v), \tag{6}$$

where σ_{v} denotes volumetric total stress and K_{dr} is the drainedisothermal bulk modulus.

The flow equations are discretized using a finite volume formulation, with solutions computed for each finite volume block/cell. The geomechanical equations are discretized using a Galerkin finite element formulation. The numerical implementation of the fixed-stress iterative scheme follows the description in Fu et al. (2021). In this work, all high-fidelity simulations (HFS), by which we mean the numerical solution of Eqs. (1), (3) and (5), are performed using GEOS, a

high-performance computing simulation environment for geoscience applications (Ju et al., 2020; Settgast et al., 2017).

2.2. Variables and domains in HFS and surrogate model

We can represent the HFS described above as

$$\left[\mathbf{p}_{f},\mathbf{S}_{f},\mathbf{d}_{f}\right]=f\left(\mathbf{m}_{f}\right),\tag{7}$$

where f denotes the numerical simulation, $\mathbf{m}_f \in \mathbb{R}^{n_b}$ is the specified geomodel, and $\mathbf{p}_f \in \mathbb{R}^{n_b \times n_{ts}}$, $\mathbf{S}_f \in \mathbb{R}^{n_b \times n_{ts}}$ and $\mathbf{d}_f \in \mathbb{R}^{3 \times n_d \times n_{ts}}$ represent the dynamic gaseous-phase pressure, gaseous-phase saturation and displacement fields (the displacement field has a component in each coordinate direction). Here n_b is the total number of finite volume grid blocks in the full-order overall model, n_d is the total number of finite element nodes in the full-order model, and n_{ts} is the number of time steps in the HFS. Note that \mathbf{p}_f , \mathbf{S}_f and \mathbf{d}_f are the primary variables computed in the HFS (i.e., all other dependent quantities can be determined from these variables).

In the numerical setup for coupled flow and geomechanics problems in carbon storage simulations, the overall model domain is much larger than the storage aquifer domain. This is necessary in order to correctly simulate the geomechanical response. The problem specification used in this study, shown in Fig. 1, illustrates the setup. Here the storage aquifer domain is of size 8000 m \times 8000 m \times 120 m, and is represented by 40 \times 40 \times 12 grid blocks. The full domain is much larger – 20 km \times 20 km \times 2 km, represented by 60 \times 60 \times 37 grid blocks. The overburden rock (red region in Fig. 1), bedrock (green region), and the surrounding domain (blue region) are the additional domains required for the geomechanical solution.

When data assimilation is performed, flow simulations for a large number (e.g., $O(10^3 - 10^6)$ of candidate geomodels \mathbf{m}_f are required. In this work we consider the porosity and (isotropic) permeability fields in the storage aquifer to be uncertain, and we fix all other geomodel parameters. This includes porosity, permeability, Young's modulus and Poisson's ratio in the surrounding domain, overburden and bedrock, and Young's modulus and Poisson's ratio in the storage aquifer. These properties could also be varied if necessary, but we would then require more training simulations for the surrogate model. In addition, most of these properties are expected to have a relatively small impact on the key quantities of interest, such as plume location and pressure buildup in the storage aquifer.

Consistent with the discussion above, new storage-aquifer geomodels are denoted $\mathbf{m}_s \in \mathbb{R}^{n_s}$, and the corresponding saturation and pressure states are defined as $\mathbf{S}_s \in \mathbb{R}^{n_s \times n_{ts}}$ and $\mathbf{p}_s \in \mathbb{R}^{n_s \times n_{ts}}$, where n_s indicates the number of grid blocks in the storage aquifer. Another quantity of interest is the vertical displacement map at the Earth's surface, directly above the storage aquifer. This is denoted as $\mathbf{d}_g \in \mathbb{R}^{n_g \times n_{ts}}$, where n_g is the number of surface nodes aligned with (i.e., directly above) the storage aquifer. Although displacement is computed in the x, y and z directions, \mathbf{d}_g only includes the vertical displacement, consistent with the data available from InSAR (Bürgmann et al., 2000) measurements.

The storage aquifer geomodels \mathbf{m}_s considered in this study are characterized by multi-Gaussian permeability and porosity fields. For a given geomodel, the permeability $\mathbf{k}_s \in \mathbb{R}^{n_s}$ and porosity $\phi_s \in \mathbb{R}^{n_s}$ descriptions are derived from the same multi-Gaussian field $\mathbf{m}_s \in \mathbb{R}^{n_s}$. These fields are generated using the geomodeling tool SGeMS (Remy et al., 2009). For a particular grid block *i* in the storage aquifer, the isotropic permeability $(k_s)_i$ and porosity $(\phi_s)_i$ are given by $(k_s)_i = \exp(a(m_s)_i + b)$ and $(\phi_s)_i = c(m_s)_i + d$, where *a*, *b*, *c* and *d* are specified constants. In this work we specify c = 0.05 and d = 0.3. Porosity will thus be positive unless the sampled value of the standard normal variable $(m_s)_i$ is six standard deviations below the mean. If a negative porosity were obtained, it would be set to a very small but positive value. This was not necessary in our runs, however, as the minimum observed value of $(\phi_s)_i$ was about 0.07. Other modeling



Fig. 1. Illustration of the overall model (left) and storage aquifer (right). The yellow, red, blue and green regions denote the storage aquifer, overburden rock, surrounding domain, and bedrock, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

parameters are kept fixed, though some of these parameters are specified to have different values in different domains, as discussed in Section 3.1.

Despite the fact that the dynamic quantities of interest (S_s , p_s , d_g) are restricted to the storage aquifer region and the ground surface directly above the storage aquifer, the coupled problem must still be solved over the full domain (which leads to high computational expense). An advantage of the deep-learning-based surrogate model developed here is that it can be trained to predict only the quantities of interest in the domains of interest. This leads to savings in the time required for training and predictions.

Given a number of HFS training 'samples' collected from the numerical simulator, a 3D recurrent R-U-Net is trained to serve as a surrogate. We use \hat{f} to denote the recurrent R-U-Net surrogate and $\hat{\mathbf{S}}_s \in \mathbb{R}^{n_s \times n_t}$, $\hat{\mathbf{p}}_s \in \mathbb{R}^{n_s \times n_t}$ and $\hat{\mathbf{d}}_g \in \mathbb{R}^{n_g \times n_t}$ to represent the surrogate-model predictions for gaseous-phase saturation, gaseous-phase pressure, and surface displacement. Note that these quantities are provided at n_t time

steps rather than at the n_{ts} HFS time steps (with n_t typically much less than n_{ts}), as the surrogate is applied to predict the states only at key time steps. The recurrent R-U-Net surrogate model is thus expressed as

$$\left[\widehat{\mathbf{p}}_{s}, \widehat{\mathbf{S}}_{s}, \widehat{\mathbf{d}}_{g}\right] = \widehat{f}(\mathbf{m}_{s}). \tag{8}$$

For conciseness, in our discussion below we will use $\hat{\mathbf{x}}$ to denote $\hat{\mathbf{S}}_s$, $\hat{\mathbf{p}}_s$ or $\hat{\mathbf{d}}_g$, and \mathbf{x} to indicate the corresponding reference solutions generated from the HFS.

2.3. Recurrent R-U-Net surrogate model for coupled problems

The 3D recurrent R-U-Net implemented in this work is similar to the network used in our previous work for two-phase flow in the absence of geomechanical effects (Tang et al., 2021). Here we describe the overall procedure, with emphasis on the new treatments required for the coupled problem.



Fig. 2. Schematic of 3D residual U-Net (R-U-Net) used in this work. This network entails encoding and decoding nets. Feature maps F_1 to F_4 extracted in the encoding net are concatenated with the upsampled features in the decoding net to generate predictions for the states.

The residual U-Net (referred to as R-U-Net), illustrated in Fig. 2, is a key component of the recurrent R-U-Net architecture. U-Nets provide robust architectures for high-dimensional matrix-wise regression problems (Ronneberger et al., 2015). In our setting, this network is used to capture the spatial correlation between the input rock properties \mathbf{m}_s and the output dynamic states. U-Nets contain encoding and decoding nets, with the extracted feature maps \mathbf{F}_1 to \mathbf{F}_4 in the encoding net concatenated with the upsampled features in the decoding net, as shown in Fig. 2. This linkage between the encoding and decoding pathways facilitates information flow in training, and leads to improved accuracy in the predicted states (Tang et al., 2020; 2021). Residual layers (He et al., 2016) are added to process the most compressed feature \mathbf{F}_5 . The residual U-Net architecture thus described can provide high accuracy for steady-state predictions, but the network in Fig. 2 is not designed to capture temporal dynamics.

In order to capture the time evolution of the saturation, pressure and displacement states, we incorporate the R-U-Net into a convolutional long-short term memory network (convLSTM). The convLSTM network (Xingjian et al., 2015) introduces recurrency, enabling us to capture dynamics associated with high-dimensional features. The overall framework, referred to as the 3D recurrent R-U-Net, is thus able to represent the temporal evolution of the 3D (saturation and pressure) and 2D (surface displacement) fields of interest.

The architecture of the 3D recurrent R-U-Net is shown in Fig. 3. The encoding net extracts feature maps \mathbf{F}_1 to \mathbf{F}_5 . These maps incorporate effects at a range of scales, with \mathbf{F}_5 representing the most compressed/global set of features. Importantly, the convLSTM propagates only \mathbf{F}_5 , meaning this is the only feature map that varies in time. The resulting representations are denoted \mathbf{F}_5^t , $t = 1, ..., n_t$, where n_t is the number of time steps at which the surrogate model provides predictions (recall $n_t < n_{ts}$). The \mathbf{F}_5^t maps are then combined with the features \mathbf{F}_1 to \mathbf{F}_4 and upsampled to provide the predicted state maps $\mathbf{\hat{x}}^t$, $t = 1, ..., n_t$. The use of the same \mathbf{F}_1 to \mathbf{F}_4 at all time steps leads to a reduction in the number of tunable network parameters and thus to savings in training time.

The performance of the recurrent R-U-Net described above is relatively insensitive to small changes in its architecture. Adding or removing one layer in the encoding net (and in the corresponding decoding net), for example, does not affect the performance significantly. The use of additional layers may lead to a slight improvement in prediction accuracy, though this increases the number of trainable parameters and thus the computational cost. For the problems considered here and in our previous studies, we have found that four encoding layers (and four corresponding decoding layers) provide an appropriate balance between prediction accuracy and computational efficiency.

As discussed in Section 2.2, the domain for the state predictions $\hat{\mathbf{x}}^t$ need not coincide precisely with that from the HFS. Thus, rather than predict vertical displacement at the n_g surface nodes lying above the storage aquifer, we predict this quantity at the tops of the blocks in the uppermost layer of the overburden domain. These displacements (for the HFS), denoted $\mathbf{d}_{gb} \in \mathbb{R}^{n_{gb} \times n_t}$, are determined from $\mathbf{d}_g \in \mathbb{R}^{n_g \times n_t}$ through a simple averaging. More specifically, the block-top displacement value is computed as the average of the four nodal values at the top surface. We thus have $n_{gb} = n_g - n_{dx} - n_{dy} + 1$, where n_{dx} and n_{dy} are the number of nodes in the *x* and *y* directions at the top surface (above the storage aquifer), with $n_g = n_{dx} \times n_{dy}$. Note we need these maps at only n_t (rather than n_{ts}) time steps.

Our prediction of $\hat{\mathbf{d}}_{gb} \in \mathbb{R}^{n_{gb} \times n_t}$ is accomplished using the same recurrent R-U-Net architecture that is used for saturation and pressure. This means that we actually predict a 3D field, and then project it to the target 2D map $\hat{\mathbf{d}}_{gb}^t$ at each time step. The recurrent R-U-Net could, in principle, be modified to predict the 2D maps directly, but this would require additional exploration of network architecture and associated hyperparameters. Our approach entails copying and concatenating the surface deformation map $\mathbf{d}_{gb}^t n_l$ times, where n_l is the number of layers in the storage aquifer. With this treatment, the displacement data are of the same dimension as \mathbf{S}_s^t and \mathbf{p}_s^t . The requisite 2D map $\hat{\mathbf{d}}_{gb}^t$ is then con-



Fig. 3. Schematic of the recurrent R-U-Net architecture. The convLSTM network accepts the global feature map \mathbf{F}_5 from the encoding net and generates a sequence of feature maps \mathbf{F}_5 , $t = 1, ..., n_t$. These are decoded, separately, into a sequence of predictions for the states $\hat{\mathbf{x}}^t$, $t = 1, ..., n_t$, using the same decoding net. Here $\hat{\mathbf{x}}$ represents the predicted states $\hat{\mathbf{S}}_s$, $\hat{\mathbf{p}}_s$ and $\hat{\mathbf{d}}_g$. Figure modified from Tang et al. (2021).

structed as the average of the result over the n_l layers. As we will see, this approach provides accurate predictions, while maintaining consistency in the network for all of the predicted quantities.

For full details on the recurrent R-U-Net architecture used in this study, please see Appendix B in Tang et al. (2021). A key advantage of this network is its flexibility in terms of reuse for different problems. This includes the ability to predict 2D maps along with 3D fields, which enables the convenient handling of surface displacement described above.

2.4. Data preprocessing and network training

A set of 2000 high-fidelity simulations of the coupled problem, with each run corresponding to a different storage-aquifer realization \mathbf{m}_{s} , is performed to provide the training data. The output state maps at prescribed time steps are saved. The boundary conditions, along with the well locations and CO₂ injection-rate specifications, are the same in all runs.

Appropriate data preprocessing is necessary to achieve optimal network performance. In our previous work (Tang et al., 2020; 2021), we considered binary channelized geomodels, which were characterized by $m_i = 0$ (background mud) and $m_i = 1$ (sand). Thus no additional scaling of the geomodel was required. Saturation values fall between 0 and 1, so no processing was required for these fields. Pressure fields were normalized through use of 'detrending' and 'min-max' scaling, which were found to provide better results than time-independent normalization.

In this study, the input multi-Gaussian fields \mathbf{m}_s , which are used to construct logk and ϕ , are standard normal variables and are thus O(1). Thus no preprocessing is applied. Saturation fields \mathbf{S}_s , again between 0 and 1, can also be used directly. For the pressure and surface displacement fields, we apply detrending and standardization, which we found to provide slightly better accuracy than our previous treatments (Tang et al., 2020; 2021). The specific preprocessing for pressure and surface displacement is given by

$$\widetilde{\mathbf{x}}_{i}^{t} = \frac{\mathbf{x}_{i}^{t} - \operatorname{mean}\left(\left[\mathbf{x}_{1}^{t}, \dots, \mathbf{x}_{n_{smp}}^{t}\right]\right)}{\operatorname{std}\left(\left[\mathbf{x}_{1}^{t}, \dots, \mathbf{x}_{n_{smp}}^{t}\right]\right)}, \quad i = 1, \dots, n_{smp}, \quad t = 1, \dots, n_{t},$$
(9)

where n_{smp} is the number of training samples. With this approach, each pressure and displacement field at time step t, \mathbf{x}_{t}^{t} , is normalized by subtracting the mean field (over all n_{smp} samples) and dividing each element by its standard deviation at that time step.

The recurrent R-U-Net weights, denoted by θ , are then optimized to minimize the difference between the direct neural network predictions $\tilde{\mathbf{x}}^t$ (before inverse transform) and the normalized HFS reference $\tilde{\mathbf{x}}^t$. This optimization can be expressed as

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \frac{1}{n_{smp}} \frac{1}{n_t} \sum_{i=1}^{n_{smp}} \sum_{t=1}^{n_t} \| \widetilde{\mathbf{x}}_i^t - \widetilde{\mathbf{x}}_i^t \|_2^2, \tag{10}$$

where θ^* indicates the optimized neural network parameters. Separate θ^* are determined for the saturation, pressure and displacement fields (i. e., we train three networks). Note that the mismatch in all quantities is measured in the L^2 norm, in contrast to our approach in Tang et al. (2021).

The neural network parameters are determined using the adaptive moment estimation (ADAM) optimizer (Kingma and Ba, 2014). The batch size is set to 4, and a total of 300 epochs are used to assure convergence. Each of the three trainings converges within 5 h on a Nvidia Tesla V100 GPU. These trainings can be performed in parallel on a single GPU given the small batch size.

3. Surrogate model evaluation

In this section, we first describe the setup for the carbon storage problem with coupled flow and geomechanics used in our evaluations. Results for 3D saturation and pressure fields and 2D surface displacement maps will then be presented. Error statistics and percentile results (P_{10} , P_{50} , P_{90}) for key quantities will also be provided.

3.1. Coupled problem setup

As discussed in Section 2 and shown in Fig. 1, the overall problem corresponds to a domain of dimensions $20 \text{ km} \times 20 \text{ km} \times 2 \text{ km}$ (in the *x*, *y* and *z* directions). The overall model is defined on a $60 \times 60 \times 37$ grid. The grid-block size is constant within the storage aquifer region (corresponding to $40 \times 40 \times 12$ blocks), but it increases in the surrounding region as we move outwards. The overburden rock and bedrock are of thickness 1380 m and 500 m respectively.

The model contains four injection wells, which are open to flow in all 12 layers in the storage aquifer. Each well injects fully-saturated supercritical CO₂ at a constant mass flow rate of 36 kg CO₂ per second. This is equivalent to 1.135×10^9 kg CO₂ per year, or 1.135 Mta (megatonnes per anum), per well. Thus the total injection into the system is 4.541 Mta, which corresponds to a large-scale storage operation. The top of the overburden rock is subject to atmospheric pressure. The simulation time frame is 30 years.

In this study, pressures at the lateral boundaries of the 20 km \times 20 km overall model are fixed at their initial values. In previous studies, both fixed-pressure and no-flow boundary conditions have been applied. The impact of different boundary specifications has been considered in several studies, including Birkholzer et al. (2015); Deng et al. (2012) and Zhang et al. (2016). The setup used here shares some similarities with that in Deng et al. (2012), where a simulation domain with lateral extent 16 km \times 16 km and fixed-pressure boundaries was considered. A range of domain sizes and boundary conditions were simulated in Zhang et al. (2016). One of the models considered was of lateral extent 10 km \times 10 km, again with fixed-pressure lateral boundary conditions. Thus our setup is consistent with previous models presented in the literature. It should be noted that in our setup (as well as in those discussed above), the pressure boundary conditions are 'felt' within the storage domain during the injection period. In practice, the locations of the boundaries and the pressure values specified should be consistent with the regional setting, pressure effects from nearby CCUS operations, brine/pressure management schemes, etc. We note finally that we do not expect the performance of our surrogate model to be strongly affected by these boundary specifications.

As noted in Section 2, the storage aquifer is characterized by multi-Gaussian fields generated using SGeMS (Remy et al., 2009). Three realizations of the resulting log-permeability fields are displayed in Fig. 4. An exponential variogram model was used, and the correlation lengths were specified to be 20 grid blocks in the *x* and *y* directions and 5 grid blocks in the *z* direction. The mean and standard deviation of log-permeability are 2.5 and 1 respectively, which leads to an (arithmetic) average permeability of 23 md in the storage aquifer. Porosity is computed (block by block) directly from log-permeability, with a mean of 0.3 and a standard deviation of 0.05. Permeability and porosity are conditioned to well data, meaning all realizations display the same properties in well blocks.

Relevant parameters for the four domains are summarized in Table 1. Note that Poisson's ratio is set to 0.2 and Biot's coefficient to 1.0 for all blocks in the overall model. As is evident from the table, the overburden and bedrock are essentially impermeable ($k = 10^{-7}$ md), while the surrounding region is characterized by k = 300 md. Porosity is set to 0.3 in overburden rock, bedrock and surrounding regions (the porosity values in the impermeable overburden and bedrock have virtually no effect on the solution). Young's modulus is set to 5 GPa in the storage aquifer, 50 GPa in the bedrock and 1 GPa in the overburden. The



Fig. 4. Multi-Gaussian log-permeability realizations honoring hard data at the four injection wells (indicated by white cylinders). The realization in (c) is the 'true' model used for data assimilation in Section 4.

Table 1

Parameters u	ised in t	the coupl	led models.
--------------	-----------	-----------	-------------

Storage aquifer parameters	Value			
Thickness	120 m			
Permeability	Multi-Gaussian field			
Porosity	Multi-Gaussian field			
Young's modulus	5 GPa			
Overburden parameters	Value			
Thickness	1380 m			
Permeability	10 ⁻⁷ md			
Porosity	0.3			
Young's modulus	1 GPa			
Bedrock parameters	Value			
Thickness	500 m			
Permeability	10 ⁻⁷ md			
Porosity	0.3			
Young's modulus	50 GPa			
Surrounding region parameters	Value			
Permeability	300 md			
Porosity	0.3			
Capillary pressure – Van Genuchten model	Value			
λ	0.254			
S_{ar}	0.11			
P_{max}	12,500 Pa			
Relative permeability – Corey model	Value			
S _{ar}	0.11			
S _{gr}	0.01			
n _a	4			
n _g	2			

bedrock serves as a rigid boundary and thus has a high Young's modulus value, while the overburden is susceptible to deformation in our model, and is therefore characterized by a low Young's modulus.

The fluid properties for CO₂ and water are obtained from correlations developed by Altunin et al. (Altunin, 1968) and steam table equations according to IAPWAS-IF97 (Wagner and Kretzschmar, 2008), respectively. At a pressure of 136 bar (corresponding to the top of the storage aquifer, at a depth of 1380 m), CO₂ viscosity and density are 0.035 cp and 484 kg/m³, and water viscosity and density are 0.43 cp and 986 kg/m³. These properties correspond to those for CO₂ and water at a temperature of 65 °C (and pressure of 136 bar). This aquifer temperature, at a depth of 1380 m, corresponds to a surface temperature of 23.6 °C and a geothermal gradient of 30 °C/km. This gradient is that used in Preisig and Prevost (2011) in a study of In Salah, where the aquifer temperature was 90 °C at a depth of nearly 2 km. Capillary pressure and relative permeability parameters are provided in Table 1. Here, λ is the exponent coefficient in the Van Genuchten capillary pressure model, S_{ar} is the irreducible aqueous-phase saturation

(required for the capillary pressure and relative permeability models), and P_{max} is the maximum capillary pressure. The quantity S_{gr} is the gaseous-phase residual saturation, and n_a and n_g are the exponents in the Corey relative permeability model. In the GEOS runs, both the phase mobilities and the absolute permeabilities were upwinded in the flow equations. This treatment leads to improved robustness, though it acts to 'mix' properties at domain boundaries.

Before presenting surrogate model predictions, we consider the sensitivity of the high-fidelity model predictions to the permeability values used for the overburden and bedrock. In this assessment, we assign these permeabilities to be $k_{over} = 10^{-5}$ md, $k_{over} = 10^{-4}$ md, and $k_{\rm over} = 10^{-3}$ md (in all cases we take bedrock permeability equal to overburden permeability k_{over}). The baseline value ($k_{over} = 10^{-7} \text{ md}$) corresponds to intact impermeable rock, while the higher values are intended to represent impermeable rock with varying amounts of discontinuities (such as small-scale fractures). The storage aquifer geomodel used in this assessment is that shown in Fig. 4c, and all other model settings are as defined above. Results for average displacement at the Earth's surface are presented in Fig. 5. This quantity is computed as the average at the surface above the four observation locations O1-O4 (defined later) and the center point of the top layer. We see in Fig. 5 that, over the range $k_{\rm over} = 10^{-7}$ md to $k_{\rm over} = 10^{-4}$ md, variation in the overburden and bedrock permeability value has little effect on the displacement results of interest. Results are impacted, however, for $k_{\rm over} = 10^{-3}$ md. Thus we conclude that, for this model, displacement



Fig. 5. Average vertical displacement at Earth's surface for different values of overburden and bedrock permeability (denoted k_{over}). Storage aquifer geomodel is that shown in Fig. 4c. Results are from high-fidelity simulator GEOS. Baseline results and those for $k_{over} = 10^{-5}$ md overlap.

results are not sensitive to the value used for permeability in the overburden and bedrock until we exceed $k_{over} = 10^{-4}$ md.

3.2. Recurrent R-U-Net saturation, pressure and displacement fields

We generate a total of 2000 random realizations of the multi-Gaussian storage-aquifer permeability and porosity fields using SGeMS. These fields are 'inserted' into the overall model description (which also includes the other three domains), and high-fidelity numerical simulation is then performed for each model. All runs are conducted using the GEOS simulator (Settgast et al., 2017). The 3D saturation and pressure fields in the storage aquifer, and the 2D surface displacement maps, are saved at 10 time steps (0.5, 2, 5, 8, 11, 15, 19, 23, 27, 30 years). These solution states constitute the training data.

A new set of 500 geomodels is generated and simulated (at high fidelity). These results represent the test cases used to assess the performance of the surrogate model. Saturation fields within the storage aquifer for a particular test case, at three different time steps, are shown in Fig. 6. For visual clarity, only grid blocks with CO₂ saturation greater than 0.01 are displayed (note that we use the term CO₂ saturation to refer to gaseous-phase saturation). The saturation error for this case (quantified later in Eq. (11)) is greater than the median error over the 500 test cases, so these results can be considered to be representative. The upper row in Fig. 6 shows the recurrent R-U-Net surrogate predictions, while the lower row displays the reference HFS results. We observe variability in the CO2 saturation distributions around the different injection wells, despite the fact that they all inject at the same rate. There is, nonetheless, very close agreement between the surrogate and HFS results. It is evident from Fig. 6c and f that some of the injected CO2 has left the storage aquifer (and entered the surrounding region, which is not shown in the figure).

In terms of timing, it takes about 0.8 h of parallel computation (on 32 cores) to complete a single high-fidelity simulation. The trained surrogate, by contrast, can provide batch-wise predictions for 100 geomodels in less than 1 s, or ≤ 0.01 s per realization. Thus the recurrent R-U-Net surrogate does indeed provide dramatic speedups for this coupled problem.

Saturation fields at the end of the 30-year simulation time frame, for

three different realizations, are presented in Fig. 7. The errors for these models are all around or above the median error for the full set of test cases. Consistent with Fig. 6, we again observe clear agreement between the HFS and surrogate model results. There is a reasonable amount of variation from case to case, which is evident upon close inspection. For example, the top of the plume from the front-most injector is much larger in Fig. 7c than in Fig. 7a, and the top of the plume from the right-most injector is larger in Fig. 7b than in Fig. 7a or c. The 3D shapes of the plumes also vary from realization to realization. The (sometimes subtle) variations between realizations are well captured by the recurrent R-U-Net model.

Results for the 3D pressure field in the storage aquifer, for a representative test case (with error above the median), are shown in Fig. 8. Because we have nonzero capillary pressure in this problem, pressure is different in the aqueous and gaseous phases. In all of the pressure results in this paper, we present gaseous-phase pressures, which we refer to simply as 'pressure.' From Fig. 8 we see that pressure displays higher local variation at early time (2 years) than at later times. Pressure does not continuously increase with time because of the very large surrounding region and pressure boundary conditions in the overall model. The correspondence between the surrogate model and the reference HFS results is again very close.

Pressure fields at 30 years, for three different test-case realizations (with error near or above the median test-case error), are displayed in Fig. 9. We see substantial differences from case to case in near-well pressure buildup, with Realization 2 displaying high pressure around wells 11 and 14. Despite the variability between and within realizations, consistent agreement between the surrogate and HFS results is again observed.

As discussed in Section 2.2, vertical displacements for the grid blocks lying directly above the storage aquifer, at the Earth's surface, are computed from the (nodal) finite element geomechanics solution. These are the quantities the surrogate model is trained to predict. Results for this ground-level vertical displacement are shown in Fig. 10. These results are for the same realization (and time steps) as in Fig. 8 (which displayed pressure fields). The upper and lower rows again show the surrogate and HFS results, and we observe a high level of accuracy in the surrogate model predictions. The surface displacements reach a



Fig. 6. CO2 saturation fields from recurrent R-U-Net surrogate model (upper row) and HFS (lower row) for a representative test case at three different times.

International Journal of Greenhouse Gas Control 118 (2022) 103692



Fig. 8. Pressure fields from recurrent R-U-Net surrogate model (upper row) and HFS (lower row) for a representative test case at three different times.

maximum of 14 cm, which is a substantial amount of deformation. Of particular interest is the correspondence between the displacement and pressure fields, evident by comparing Figs. 8 and 10. Specifically, the maximum pressures are observed near wells I1 and I4, and it is above these wells that the maximum displacements occur.

Surface displacement results for the three test-case models presented in Fig. 9 (for pressure) are shown in Fig. 11. These results are all at a time of 30 years. Despite the significant differences between the various cases, we continue to observe close correspondence between the surrogate model and HFS results. The highest near-well pressures are observed for well I4 in Realization 2, and the largest vertical displacements are also seen above this well for this realization. Analogously, lower near-well pressures and surface displacements are evident for Realization 1.

3.3. Surrogate model errors and percentile predictions

We now quantify the saturation, pressure and surface displacement errors associated with the surrogate model predictions. We compute relative errors for these three quantities, defined as

International Journal of Greenhouse Gas Control 118 (2022) 103692



Fig. 10. Vertical displacement maps from recurrent R-U-Net surrogate model (upper row) and HFS (lower row) for a test case at three different times. These results are for the same realization as in Fig. 8.

International Journal of Greenhouse Gas Control 118 (2022) 103692



Fig. 11. Vertical displacement maps from recurrent R-U-Net surrogate model (upper row) and HFS (lower row) for three different test cases at 30 years. These results are for the same realizations as in Fig. 9.

$$\delta_{S} = \frac{1}{n_{e}n_{s}n_{t}} \sum_{i=1}^{n_{e}} \sum_{j=1}^{n_{s}} \sum_{t=1}^{n_{s}} \frac{\|(\widehat{S}_{s})_{i,j}^{t} - (S_{s})_{i,j}^{t}\|}{(S_{s})_{i,j}^{t} + \epsilon},$$
(11)

$$\delta_p = \frac{1}{n_e n_s n_t} \sum_{i=1}^{n_e} \sum_{j=1}^{n_s} \sum_{t=1}^{n_s} \frac{\|\left(\hat{p}_s\right)_{i,j}^t - (p_s)_{i,j}^t\right\|}{(p_s)_{i,\max}^t - (p_s)_{i,\min}^t},$$
(12)

$$\delta_{d} = \frac{1}{n_{e}n_{gb}n_{t}} \sum_{i=1}^{n_{e}} \sum_{j=1}^{n_{gb}} \sum_{t=1}^{n_{t}} \frac{\|\left(\widehat{d}_{gb}\right)_{i,j}^{t} - \left(d_{gb}\right)_{i,j}^{t}\|}{\left(d_{gb}\right)_{i,j}^{t}},$$
(13)

where n_e is the number of test cases, n_s is the total number of grid blocks in the storage aquifer, n_t is the number of time steps considered, and n_{gb} is the number of grid blocks on the surface above the storage aquifer. A constant e = 0.01 is introduced in the denominator of Eq. (11) to avoid division by very small values. Note that the pressure error is normalized by the difference between the maximum and minimum pressure values for test case *i* at time step $t ((p_s)_{i,max}^t - (p_s)_{i,min}^t)$. This treatment leads to larger pressure errors than would be computed if we used the absolute pressure value $(p_s)_{i,i}^t$.

Applying Eqs. (11)–(13), we obtain $\delta_S = 5.3\%$, $\delta_p = 0.31\%$ and $\delta_d = 1.2\%$. These small relative errors demonstrate that the recurrent R-U-Net surrogate predicts storage aquifer saturation and pressure, along with surface-level vertical displacement, quite accurately. The observation that $\delta_S > \delta_p$ is consistent with our previous findings (Tang et al., 2020; 2021). Saturation errors are generally largest in the vicinity of fronts.

We next assess the statistical correspondence in block-wise results between the surrogate and HFS models. This evaluation enables us to see whether the surrogate model maintains accuracy over a range of possible test-case responses. Specifically, for four observation locations in the top layer of the storage aquifer (shown in Fig. 12 and denoted O1–O4), we compute the 10th, 50th and 90th percentile results for saturation and pressure as a function of time. These quantities are referred to as the P_{10} , P_{50} and P_{90} responses. Each curve can correspond to a different realization from time step to time step.

Saturation and pressure results are presented in Figs. 13 and 14. The solid black curves are the target HFS results, and the dashed red curves are the surrogate model results. The lower curve represents the P_{10} result, the middle curve is P_{50} , and the upper curve P_{90} . Agreement between the HFS and surrogate model P_{10} , P_{50} and P_{90} results, for both saturation and pressure at all four observation locations, is excellent. This suggests that the surrogate model is accurate over the full P_{10} – P_{90} range for these quantities. The consistent agreement in saturation results, despite the wide variability for this quantity (evident in the



Fig. 12. Storage aquifer realization showing injection wells I1–I4 (white cylinders) and observation locations O1–O4 (black arrows).



Fig. 13. Comparison of saturation statistics at four observation locations over the full ensemble of 500 test cases. Red and black curves represent P₁₀ (lower), P₅₀ (middle) and P₉₀ (upper) results from the recurrent R-U-Net surrogate model and the high-fidelity simulator, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

difference between the P_{10} and P_{90} curves in Fig. 13), is particularly noteworthy.

For surface displacement, the observation locations are at ground level, directly above the O1–O4 storage-aquifer locations. The P₁₀, P₅₀ and P_{90} results at the four observation locations are shown in Fig. 15. Agreement between HFS and surrogate predictions is again very close, though there is relatively little variation between the P₁₀ and P₉₀ surface displacement curves. Nonetheless, as we will see in Section 4, data of this type can still be used in history matching to reduce prediction uncertainty.

4. Data assimilation using recurrent R-U-Net surrogate model

In this section we apply the recurrent R-U-Net surrogate model for data assimilation. A rigorous rejection sampling (RS) procedure is applied to provide posterior (history matched) samples conditioned to surface displacement data. We first describe the RS procedure, and then present data assimilation results.

4.1. Rejection sampling using PCA-based geomodels

RS is a formal posterior sampling method that can correctly quantify posterior uncertainty. It has the drawback of requiring very large numbers of samples and function evaluations, which renders it impractical in most settings when HFS is used for the function evaluations. With the surrogate model applied in place of the high-fidelity simulator, however, RS can be applied in a much wider range of cases.

In our example, we construct and evaluate 500,000 realizations of the geomodel. Rather than apply the geological modeling software SGeMS to generate each realization, we construct these models very quickly using principal component analysis (PCA). Following the description in Liu and Durlofsky (2021), we construct the PCA basis

matrix, in a preprocessing step, from a set of 3000 SGeMS models. Then, new PCA realizations $\mathbf{m}_{s}^{\text{pca}} \in \mathbb{R}^{n_{s}}$ are constructed through application of $\mathbf{m}_{s}^{\text{pca}} = \mathbf{\Phi} \boldsymbol{\xi} + \overline{\mathbf{m}}_{s}$, where $\mathbf{\Phi} \in \mathbb{R}^{n_{s} \times l}$ is the basis matrix, $\overline{\mathbf{m}}_{s} \in \mathbb{R}^{n_{s}}$ is the mean of the SGeMS models, and $\boldsymbol{\xi} \in \mathbb{R}^l$ is a random variable, with each component sampled independently from $\mathcal{N}(0,1)$. The dimension of $\boldsymbol{\xi}(l)$ is set to 2000, which acts to retain 95% of the total 'energy' in the original set of 3000 realizations. This provides \mathbf{m}_{s}^{pca} realizations that are essentially indistinguishable visually from SGeMS models while avoiding overfitting.

Note that, in Tang et al. (2020, 2021), a convolutional neural network (CNN)-based parameterization technique (called CNN-PCA) was used for realization generation instead of standalone PCA. CNN-PCA was required in those studies because the geomodels were channelized, and standalone PCA is not applicable for such systems. We can, however, directly apply PCA here since the geomodels are multi-Gaussian.

We now present the RS procedure applied in this work. This description follows that in Tang et al. (2021), though here we use PCA instead of CNN-PCA, and the observed data are surface displacements rather than flow rates at wells.

- Sample each component of $\boldsymbol{\xi} \in \mathbb{R}^l$ from $\mathcal{N}(0,1)$. Construct $\mathbf{m}_s^{\text{pca}}(\boldsymbol{\xi})$ as described above.
- Sample a probability *p* from a uniform distribution in [0,1].
- Compute the likelihood function $L(\mathbf{m}_s^{\text{pca}}(\boldsymbol{\xi}))$ using

$$L(\mathbf{m}_{s}^{\text{pca}}(\boldsymbol{\xi})) = c \exp\left(-\frac{1}{2}\left[\widehat{f}\left(\mathbf{m}_{s}^{\text{pca}}(\boldsymbol{\xi})\right) - \mathbf{d}_{\text{obs}}\right]^{T} C_{D}^{-1}\left[\widehat{f}\left(\mathbf{m}_{s}^{\text{pca}}(\boldsymbol{\xi})\right) - \mathbf{d}_{\text{obs}}\right]\right),$$
(14)

where *c* is a normalization constant, $\hat{f}(\mathbf{m}_{s}^{\text{pca}}(\boldsymbol{\xi}))$ indicates the surrogate model predictions for surface displacement for geomodel



(c) Pressure at O3

(d) Pressure at O4

Fig. 14. Comparison of pressure statistics at four observation locations over the full ensemble of 500 test cases. Red and black curves represent P_{10} (lower), P_{50} (middle) and P_{90} (upper) results from the recurrent R-U-Net surrogate model and the high-fidelity simulator, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

 $\mathbf{m}_{s}^{pca}(\boldsymbol{\xi})$, \mathbf{d}_{obs} denotes the observed surface displacement data, and C_{D} is the covariance matrix of data measurement error.

• Accept $\mathbf{m}_{s}^{\text{pca}}(\boldsymbol{\xi})$ if $p \leq \frac{L(\mathbf{m}_{\text{pca}}(\boldsymbol{\xi}))}{S_{L}}$, where S_{L} is the maximum likelihood value over all prior models considered. Otherwise, reject $\mathbf{m}_{s}^{\text{pca}}(\boldsymbol{\xi})$.

As is evident from the description above, a randomly generated realization $m_s^{pca}(\xi)$ is much more likely to be accepted as a posterior sample if the predicted surface displacements are close to d_{obs} . Similarly, if the mismatch is large, the model has a very low probability of acceptance.

4.2. Problem setup and rejection sampling results

A randomly selected SGeMS gemodel, which is shown in Fig. 4(c), is specified to be the 'true' model. High-fidelity simulation is performed on this model, and the results are taken as the 'true' data \mathbf{d}_{true} . The observed data \mathbf{d}_{obs} are obtained by randomly perturbing \mathbf{d}_{true} with measurement error $\boldsymbol{\epsilon}$

$$\mathbf{d}_{\rm obs} = \mathbf{d}_{\rm true} + \boldsymbol{\epsilon},\tag{15}$$

where ϵ is sampled with **0** mean and covariance C_D . In this study the observed data correspond to surface displacement data at four observation locations (above O1–O4) at two different times (5 years and 8 years after the start of CO₂ injection). Thus we have a total of eight measurements. The standard deviation of the measurement error is set to 5% of the corresponding true data.

Using a single Nvidia Tesla V100 GPU, it takes less than 1 h for the surrogate model to provide the saturation, pressure and surface displacement predictions for the 500,000 geomodels evaluated during the RS procedure. The time required to generate the PCA realizations is negligible. High-fidelity numerical simulation, however, requires about

0.8 h per run (using 32 CPU cores). Therefore, the total computation time to perform this RS assessment using HFS would be about 400,000 h, which is clearly impractical.

With the setup described above, RS accepts 601 models out of the 500,000 considered. Data assimilation results for (ground-level) surface displacement are shown in Fig. 16. In this figure, the gray regions indicate the prior P_{10} – P_{90} intervals, the red circles denote the observed data, and the red curves indicate the surface displacement response for the true model (the observed and true data deviate slightly due to measurement error ϵ). The blue dashed curves depict the P_{10} (lower), P_{50} (middle) and P_{90} (upper) posterior RS results using the surrogate model. We also perform high-fidelity simulations on the 601 accepted geomodels. The resulting P_{10} , P_{50} and P_{90} curves are indicated by the solid black curves in Fig. 16.

The results in Fig. 16 demonstrate uncertainty reduction in surface displacement predictions at the four observation locations. The prior uncertainty above location O3 is relatively small, however, and is only modestly reduced by the data assimilation procedure. We observe close correspondence between the P_{10} , P_{50} and P_{90} predictions from the surrogate model (blue dashed curves) and those from the high-fidelity simulations (solid black curves). This is an important consistency, and suggests that the surrogate model is indeed appropriate for the many function evaluations required by RS.

Of particular interest in carbon storage operations are the pressure buildup at the top of the storage formation (caprock) and the location of the CO_2 plume. We now assess the degree to which uncertainty in these quantities is reduced through assimilation of surface displacement data. We reiterate that no data other than the eight surface displacement measurements are used for history matching. Results for prior and posterior pressure and saturation data at locations O1–O4 are shown in Figs. 17 and 18. The surface displacement data are clearly informative in terms of pressure buildup, as we see substantial uncertainty reduction at



(c) Displacement at the surface above O3

(d) Displacement at the surface above O4

Fig. 15. Comparison of vertical displacement statistics at four observation locations over the full ensemble of 500 test cases. Red and black curves represent P_{10} (lower), P_{50} (middle) and P_{90} (upper) results from the recurrent R-U-Net surrogate model and the high-fidelity simulator, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

all locations except O3 (consistent with Fig. 16). The surface displacement data are not very informative for saturation, and we see little uncertainty reduction in Fig. 18. We do, however, observe that the posterior distributions for saturation at locations O2 and O3 have shifted towards the true result (red curve), and that the true response is captured within the posterior P_{10} – P_{90} interval. Finally, consistent with our observations in Fig. 16, the surrogate and HFS P_{10} , P_{50} and P_{90} curves continue to agree closely in Figs. 17 and 18.

Finally, in Fig. 19, we present surface displacement maps for prior and posterior models. The particular models displayed correspond to the P_{10} , P_{50} and P_{90} responses for average vertical displacement, with the average computed from the surface displacements above the four observation locations, at 30 years. In the upper row, which corresponds to prior results, we observe significant discrepancies between the three maps. The three maps on the lower row, however, are much more similar to one another. This clearly illustrates the reduction in uncertainty achieved by assimilating the surface displacement data.

5. Concluding remarks

In this study, we extended our 3D recurrent R-U-Net surrogate model, originally developed for two-phase subsurface flow, to treat systems involving coupled flow and geomechanics. The methodology was then applied for the simulation of CO₂ storage operations. The recurrent R-U-Net involves the use of residual U-Nets and a convolutional long-short term memory network. The 3D solution domain for the full-order problem includes a storage aquifer, a large surrounding region, overburden, and bedrock. The storage-aquifer geomodel is viewed as uncertain, with grid-block porosity and log-permeability characterized by multi-Gaussian random fields. An advantage of the surrogate model is that it can be trained to predict key quantities only in regions of interest. In this work the recurrent R-U-Net was trained to predict CO_2 saturation (i.e., plume location) and pressure in the storage aquifer, and vertical displacement at the Earth's surface.

The problem setup entailed injection of 4.54 Mta of supercritical CO₂ via four vertical injection wells. The overall model domain contained 133,200 cells, with the storage aquifer defined by 19,200 cells. A total of 2000 full-order models, each requiring 0.8 h of simulation time on a 32-core CPU, were used for training. Following training, predictions of key quantities for new geomodels were achieved in about 0.01 s. A high degree of accuracy between the full-order HFS and the surrogate model was demonstrated for a set of 500 new test cases. Close agreement was observed for CO₂ saturation and pressure in the storage aquifer and for vertical displacement at the Earth's surface. Results were presented both for individual geomodels and for ensemble statistics (P_{10} , P_{50} , P_{90} responses over the full set of test cases) at specific observation locations.

We then applied the recurrent R-U-Net model for data assimilation. A rigorous rejection sampling algorithm was used, which required the evaluation of a large number of prior models. Specifically, 500,000 prior geomodels were generated using PCA, and flow predictions were provided using the surrogate model (this assessment would be intractable using HFS). A small number of surface displacement measurements, derived in this case from numerical simulation results on a synthetic 'true' model, were used for data assimilation. Significant uncertainty reduction in predictions for both surface displacement and storage aquifer pressure was observed.

Future work in this area should address a number of topics. Larger and more complicated models, involving uncertainty in a wider range of properties including geomechanical parameters, should be considered. The use of other data assimilation algorithms, with a variety of data types, is also of interest. Because function evaluations with the surrogate model are so fast, methods and assessments (e.g., sensitivity of



(c) Displacement at the surface above O3

(d) Displacement at the surface above O4

Fig. 16. History matching results for surface displacement at four observation locations. Gray regions represent the prior P_{10} – P_{90} range, red points and red lines denote observed and true data, solid black and dashed blue curves denote the posterior P_{10} (lower), P_{50} (middle) and P_{90} (upper) predictions obtained using the high-fidelity simulator (black) and the surrogate (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



Fig. 17. History matching results for pressure at observation locations O1–O4. Note that only surface displacement data are used in the data assimilation. Gray regions represent the prior P_{10} – P_{90} range, red lines denote true data, solid black and dashed blue curves denote the posterior P_{10} (lower), P_{50} (middle) and P_{90} (upper) predictions obtained using the high-fidelity simulator (black) and the surrogate (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

12

10

- 8

- 6

4

12

10

8

6 4



Fig. 19. Prior (upper row) and posterior (lower row) P10, P50 and P90 surface displacement maps from the surrogate model. Percentiles are based on average displacement at the surface, with the averages over locations directly above O1-O4, at 30 years.

predictions to different data types or data precision) that would be impractical with full-order numerical models can now be considered. Finally, the use of the general recurrent R-U-Net framework for other coupled multi-physics problems, involving, e.g., some combination of flow, geomechanics, thermal effects and chemical reactions, should be explored.

Code availability

The code used in this study is available at https://github.com/solo ist96/Recurrent-R-U-Net-for-CO2-Storage.git.

CRediT authorship contribution statement

Meng Tang: Conceptualization, Methodology, Software, Writing – original draft. **Xin Ju:** Data curation, Visualization. **Louis J. Durlofsky:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are grateful to the Stanford Smart Fields Consortium and to Stanford–Chevron CoRE for partial funding of this work. We thank Pengcheng Fu from the Computational Geosciences Group at Lawrence Livermore National Laboratory for his help with the construction of the coupled flow-geomechanics models and Yifu Han at Stanford for useful discussions. We also acknowledge the Livermore Computing Center and the Stanford Center for Computational Earth & Environmental Science (CEES) for providing the computational resources used in this work.

References

- International Energy Agency, 2020. Energy Technology Perspectives 2020 Special Report on Carbon Capture Utilisation and Storage, CCUS in Clean Energy Transitions.
- Alghamdi, A., Hesse, M., Chen, J., Villa, U., Ghattas, O., 2021. Bayesian poroelastic aquifer characterization from inSAR surface deformation data. Part II: quantifying the uncertainty. arXiv preprint arXiv:2102.04577.
- Alghamdi, A., Hesse, M.A., Chen, J., Ghattas, O., 2020. Bayesian poroelastic aquifer characterization from inSAR surface deformation data. Part I: maximum a posteriori estimate. Water Resour. Res. 56 (10).e2020WR027391
- Altunin, V.V., 1968. Thermophysical Properties of Carbon Dioxide. Wellingborough, Collets.
- Biot, M.A., 1941. General theory of three-dimensional consolidation. J. Appl. Phys. 12 (2), 155–164.
- Birkholzer, J.T., Oldenburg, C.M., Zhou, Q., 2015. CO₂ migration and pressure evolution in deep saline aquifers. Int. J. Greenh. Gas Control 40, 203–220.
- Bürgmann, R., Rosen, P.A., Fielding, E.J., 2000. Synthetic aperture radar interferometry to measure Earth's surface topography and its deformation. Annu. Rev. Earth Planet. Sci. 28 (1), 169–209.
- Cameron, D.A., Durlofsky, L.J., Benson, S.M., 2016. Use of above-zone pressure data to locate and quantify leaks during carbon storage operations. Int. J. Greenh. Gas Control 52, 32–43.
- Chen, B., Harp, D.R., Lu, Z., Pawar, R.J., 2020. Reducing uncertainty in geologic CO_2 sequestration risk assessment by assimilating monitoring data. Int. J. Greenh. Gas Control 94, 102926.
- Coussy, O., 2004. Poromechanics. John Wiley & Sons.
- Deng, H., Stauffer, P.H., Dai, Z., Jiao, Z., Surdam, R.C., 2012. Simulation of industrialscale CO₂ storage: multi-scale heterogeneity and its impacts on storage capacity, injectivity and leakage. Int. J. Greenh. Gas Control 10, 397–418.
- Eshiet, K., Sheng, Y., 2014. Investigation of geomechanical responses of reservoirs induced by carbon dioxide storage. Environ. Earth Sci. 71 (9), 3999–4020.
- Florez, H., Gildin, E., 2019. Model-order reduction of coupled flow and geomechanics in ultra-low permeability ULP reservoirs. SPE Reservoir Simulation Conference. Society of Petroleum Engineers.
- Fu, P., Ju, X., Huang, J., Settgast, R.R., Liu, F., Morris, J.P., 2021. Thermo-poroelastic responses of a pressure-driven fracture in a carbon storage reservoir and the implications for injectivity and caprock integrity. Int. J. Numer. Anal. Methods Geomech. 45 (6), 719–737.

- Fuchs, S.J., Espinoza, D.N., Lopano, C.L., Akono, A.-T., Werth, C.J., 2019. Geochemical and geomechanical alteration of siliciclastic reservoir rock by supercritical CO₂saturated brine formed during geological carbon sequestration. Int. J. Greenh. Gas Control 88, 251–260.
- González-Nicolás, A., Baù, D., Alzraiee, A., 2015. Detection of potential leakage pathways from geological carbon storage by fluid pressure data assimilation. Adv. Water Resour. 86, 366–384.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.
- Jahandideh, A., Hakim-Elahi, S., Jafarpour, B., 2021. Inference of rock flow and mechanical properties from injection-induced microseismic events during geologic CO₂ storage. Int. J. Greenh. Gas Control 105, 103206.
- Jha, B., Bottazzi, F., Wojcik, R., Coccia, M., Bechor, N., McLaughlin, D., Herring, T., Hager, B.H., Mantica, S., Juanes, R., 2015. Reservoir characterization in an underground gas storage field using joint inversion of flow and geodetic data. Int. J. Numer. Anal. Methods Geomech. 39 (14), 1619–1638.
- Jiang, Z., Tahmasebi, P., Mao, Z., 2021. Deep residual u-net convolution neural networks with autoregressive strategy for fluid flow predictions in large-scale geosystems. Adv. Water. Resour. 103878.
- Jin, Z.L., Garipov, T., Volkov, O., Durlofsky, L.J., 2020. Reduced-order modeling of coupled flow and quasistatic geomechanics. SPE J. 25 (01), 326–346.
- Ju, X., Fu, P., Settegast, R., Morris, J., 2021. A simple method to simulate thermo-hydromechanical processes in leakoff-dominated hydraulic fracturing with application to geological carbon storage. Earth Space Sci. Open Arch. 53.
- Ju, X., Liu, F., Fu, P., White, M.D., Settgast, R.R., Morris, J.P., 2020. Gas production from hot water circulation through hydraulic fractures in methane hydrate-bearing sediments: THC-coupled simulation of production mechanisms. Energy Fuels 34 (4), 4448–4465.
- Jung, Y., Zhou, Q., Birkholzer, J.T., 2015. On the detection of leakage pathways in geological CO₂ storage systems using pressure monitoring data: impact of model parameter uncertainties. Adv. Water Resour. 84, 112–124.
- Kim, J., Tchelepi, H.A., Juanes, R., 2011. Stability and convergence of sequential methods for coupled flow and geomechanics: fixed-stress and fixed-strain splits. Comput. Methods Appl. Mech. Eng. 200 (13–16), 1591–1606.
- Kingma, D. P., Ba, J., 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Li, C., Laloui, L., 2016. Coupled multiphase thermo-hydro-mechanical analysis of supercritical CO_2 injection: benchmark for the In Salah surface uplift problem. Int. J. Greenh. Gas Control 51, 394–408.
- Liu, M., Grana, D., 2020. Petrophysical characterization of deep saline aquifers for CO₂ storage using ensemble smoother and deep convolutional autoencoder. Adv. Water Resour. 142, 103634.
- Liu, Y., Durlofsky, L.J., 2021. 3D CNN-PCA: a deep-learning-based parameterization for complex geomodels. Comput. Geosci. 148, 104676.
- Mo, S., Zhu, Y., Zabaras, N., Shi, X., Wu, J., 2019. Deep convolutional encoder-decoder networks for uncertainty quantification of dynamic multiphase flow in heterogeneous media. Water Resour. Res. 55 (1), 703–728.
- Preisig, M., Prevost, J.H., 2011. Coupled multi-phase thermo-poromechanical effects. case study: CO₂ injection at in Salah, Algeria. Int. J. Greenh. Gas Control 5, 1055–1064.
- Remy, N., Boucher, A., Wu, J., 2009. Applied Geostatistics with SGeMS: A User's Guide. Cambridge University Press.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.
- Rutqvist, J., Rinaldi, A.P., Vilarrasa, V., Cappa, F., 2019. Numerical geomechanics studies of geological carbon storage (GCS). Science of Carbon Storage in Deep Saline Formations. Elsevier, pp. 237–252.
- Settgast, R.R., Fu, P., Walsh, S.D.C., White, J.A., Annavarapu, C., Ryerson, F.J., 2017. A fully coupled method for massively parallel simulation of hydraulically driven fractures in 3-dimensions. Int. J. Numer. Anal. Methods Geomech. 41 (5), 627–653.
- Shi, J.-Q., Smith, J., Durucan, S., Korre, A., 2013. A coupled reservoir simulationgeomechanical modelling study of the CO₂ injection-induced ground surface uplift observed at Krechba, In Salah. Energy Procedia 37, 3719–3726.
- Talebian, M., Al-Khoury, R., Sluys, L.J., 2013. A computational model for coupled multiphysics processes of CO₂ sequestration in fractured porous media. Adv. Water Resour. 59, 238–255.
- Tang, M., 2018. History Matching Production and Displacement Data Using Derivative-Free Optimization. Stanford University. Master's thesis.
- Tang, M., Liu, Y., Durlofsky, L.J., 2020. A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. J. Comput. Phys. 413, 109456.
- Tang, M., Liu, Y., Durlofsky, L.J., 2021. Deep-learning-based surrogate flow modeling and geological parameterization for data assimilation in 3D subsurface flow. Comput. Methods Appl. Mech. Eng. 376, 113636.
- de la Torre Guzman, J., Babaei, M., Shi, J.-Q., Korre, A., Durucan, S., 2014. Coupled flowgeomechanical performance assessment of CO₂ storage sites using the ensemble Kalman filter. Energy Procedia 63, 3475–3482.
- Vilarrasa, V., Bolster, D., Olivella, S., Carrera, J., 2010. Coupled hydromechanical modeling of CO₂ sequestration in deep saline aquifers. Int. J. Greenh. Gas Control 4 (6), 910–919.
- Wagner, W., Kretzschmar, H.-J., 2008. IAPWS industrial formulation 1997 for the thermodynamic properties of water and steam. International Steam Tables: Properties of Water and Steam Based on the Industrial Formulation IAPWS-IF97, pp. 7–150.

M. Tang et al.

Wen, G., Tang, M., Benson, S.M., 2021. Towards a predictor for CO₂ plume migration using deep neural networks. Int. J. Greenh. Gas Control 105, 103223.

- Wilschut, F., Peters, E., Visser, K., Fokker, P.A., van Hooff, P., 2011. Joint history matching of well data and surface subsidence observations using the ensemble Kalman filter: a field study. SPE Reservoir Simulation Symposium. Society of Petroleum Engineers.
- Xingjian, S., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., Woo, W.-C., 2015. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. Advances in Neural Information Processing Systems, pp. 802–810.
- Zhang, L., Dilmore, R.M., Bromhal, G.S., 2016. Effect of outer boundary condition, reservoir size, and CO₂ effective permeability on pressure and CO₂ saturation predictions under carbon sequestration conditions. Greenh. Gas Sci. Technol. 5, 1–15.
- Zoccarato, C., Baù, D., Ferronato, M., Gambolati, G., Alzraiee, A., Teatini, P., 2016. Data assimilation of surface displacements to improve geomechanical parameters of gas storage reservoirs. J. Geophys. Res. 121 (3), 1441–1461.