Petroleum Science 22 (2025) 1712–1735

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Deep learning-based upscaling for CO_2 injection into saline aquifers

Yan-Ji Wang [a, b], Yan Jin [c, d, *], Bo-Tao Lin [a, c], Hui-Wen Pang [c, e]

[a] College of Artificial Intelligence, China University of Petroleum (Beijing), Beijing, 102249, China
[b] Department of Energy Science and Engineering, Stanford University, CA, 94305, USA
[c] State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[d] College of Petroleum Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[e] College of Science, China University of Petroleum (Beijing), Beijing, 102249, China

ARTICLE INFO

Article history:
Received 30 September 2024
Received in revised form
10 February 2025
Accepted 11 February 2025
Available online 14 February 2025

Edited by Yan-Hua Sun

Keywords:
Artificial intelligence
Carbon storage
Subsurface flow simulation
Upscaling
Deep learning

ABSTRACT

Numerical simulation is an essential technique for CO_2 geological storage operations. However, high-resolution geological models typically consist of a large number of grid blocks, making numerical simulations computationally expensive and time-consuming. Upscaling methods are commonly used to coarsen the fine-scale geological model, with global flow-based upscaling methods generally demonstrating the highest accuracy. However, since these methods require solving flow equations over the global domain, which is still time-consuming, their applications are typically limited to cases where the coarse model is reused repeatedly (e.g., history matching or optimization). To overcome these limitations, this study develops a novel deep learning (DL)-based upscaling framework for the simulation of CO_2 injection into saline aquifers. The framework incorporates convolutional neural networks (CNNs), Transformer encoders, and Fourier neural operators (FNOs) to construct surrogate models for upscaled well index, permeability, relative permeability, and capillary pressure. A preprocessing procedure is first applied to address the issue of inaccurate upscaled parameters, which are typically caused by weak flow conditions in traditional upscaling computations. Then the surrogate models are trained using relevant local information, and the trained surrogate models are used to replace traditional numerical upscaling computations, enabling instantaneous and parallel predictions of upscaled parameters. Two representative flow patterns (left-to-right and bottom-to-top) are considered to evaluate the framework's performance. The results demonstrate that the DL-based framework significantly improves computational efficiency, achieving a speedup factor of approximately 1133 times compared to traditional upscaling methods. Additionally, it maintains or even enhances simulation accuracy, as the surrogate models correct inaccurate upscaled parameters.

1. Introduction

Numerical simulation is essential for CO_2 geological storage operations to ensure accurate predictions of CO_2 behavior and storage capacity. However, geological models are typically generated at high resolutions to capture the significant heterogeneity that occurs across various length scales (Durlofsky, 2005; Zhang et al., 2021). Direct numerical simulations of these high-resolution geological models are usually computationally expensive. Furthermore, when uncertainty quantification, history matching or optimization is required, a substantial number of simulation runs are necessary, which further increases computational costs. To address this issue, upscaling methods are often employed to coarsen high-resolution models to accelerate the simulation process.

Researchers have developed upscaling methods for CO_2 geological storage simulation with notable works described as follows. Mouche et al. (2010) proposed a homogenization-based technique to upscale the vertical migration of a CO_2 plume using a one-dimensional vertical model filled with a periodic porous medium. Saadatpoor et al. (2011) applied average-based methods to calculate upscaled permeability and porosity for CO_2 migration models, discovering that their method smoothed the capillary pressure field, leading to potential errors. Behzadi and Alvarado

* Corresponding author.
E-mail address: jiny@cup.edu.cn (Y. Jin).

https://doi.org/10.1016/j.petsci.2025.02.003

(2012) and Bouquet et al. (2013) also employed average-based calculations to obtain equivalent coarse-scale permeability. However, these studies utilized analytical methods that are generally less accurate than the flow-based upscaling methods (Durlofsky, 2005; Efendiev and Durlofsky, 2004; Renard and De Marsily, 1997). Rabinovich et al. (2015) developed a flow-based upscaling procedure for CO_2 storage simulation, but this work primarily used global upscaling methods, which entail simulations over global domain and remain computationally expensive. Moreover, this study employed simplified models, neglecting critical factors such as solubility effects, which are essential for accurately capturing the complexities of CO_2-water dynamics.

Machine learning (including deep learning) techniques have shown increasing potential for applications in subsurface flow simulation. Tang and Durlofsky (2022), Wang N. et al. (2023a), Tang et al. (2021) employed machine learning algorithms to optimize well locations or control parameters. While Han et al. (2024), Jo et al. (2021), Li and Misra (2021), Srinivasan et al. (2021) utilized machine learning methods to accelerate history matching processes. Moreover, Wen et al. (2022, 2023) and Wang N. et al. (2023b) employed deep learning models as surrogates for numerical simulations, providing rapid predictions for subsurface flow problems. There have also been limited studies exploring machine learning in the context of upscaling, where these methods are employed to accelerate upscaling calculation processes. Bohne (2018) utilized least squares and kernel ridge regression algorithms to establish a mapping from fine-scale to coarse-scale permeability. However, the employed upscaling methods and machine learning algorithms were relatively rudimentary, leading to limited predictive accuracy. Wang Y. et al. (2022, 2023) proposed machine learning-assisted upscaling methods for water–oil systems, which employed convolutional neural networks (CNNs) and machine learning regression algorithms to predict upscaled relative permeabilities. However, these studies only involved machine learning methods for local relative permeability upscaling. Similarly, Wang N. et al. (2023b) trained CNN models to predict upscaled hydraulic conductivity but only considered simple single-phase flow scenarios.

The purpose of this paper is to leverage deep learning (DL) techniques to achieve efficient upscaling for simulating CO_2 injection into saline aquifers. We consider full-physics geological models that account for mutual solubility of CO_2 and brine, salt precipitation, relative permeability hysteresis, gravity, compressibility of rock and fluids, and capillary heterogeneity. To address these complexities, we propose a novel upscaling framework that integrates flow-based upscaling methods for transmissibility, well index, and relative permeability, as well as a steady-state upscaling method for heterogeneous capillary pressure. DL surrogate models are developed for these upscaled parameters by combining advanced architectures, including convolutional neural networks (CNNs), Transformer encoders, and Fourier neural operators (FNOs).

To our knowledge, this study represents the first application of DL techniques to upscaling in CO_2 storage simulations. By leveraging a big dataset of 8000 geological models, we demonstrate the effectiveness of the proposed approach in two representative flow patterns: left-to-right and bottom-to-top. Preprocessing steps are introduced to address the issue of inaccurate upscaled parameters, ensuring the reliability of the surrogate models. After training, the DL predictions replace traditional numerical upscaling computations, providing not only significantly improved efficiency but also enhanced accuracy and robustness. This approach enables the instantaneous and parallel prediction of upscaled parameters with improved accuracy, making it a practical solution for CO_2 storage simulations.

The remainder of this paper is organized as follows. In Section 2, we present the governing equation of the full-physics geological model and describe the upscaling methodology. Section 3 details the generation of geological models. In Section 4, we develop the deep learning surrogate models to predict upscaled parameters. The overall workflow is introduced in Section 5. The results are presented in Section 6. In Section 7, we discuss the findings, and finally, conclusions and future work are provided in Section 8.

2. Governing equation and upscaling methodology

2.1. Governing equation

We consider a multiphase flow system that includes CO_2 (non-wetting phase) and brine (wetting phase). This system accounts for the mutual solubility of CO_2 and brine, as well as salt precipitation, but excludes molecular diffusion and mineralization effects since they are not significant during the CO_2 injection stage. The general form for the mass accumulation of component α, which could be either CO_2 or brine, is defined as follows:

$$\frac{\partial\left(\phi\sum_j S_j\rho_j X_j^\alpha\right)}{\partial t}+\nabla\cdot F^\alpha=q^\alpha, \tag{1}$$

where α represents a component; t is the time; q^α is the volumetric source of component α; ϕ is the porosity; S_j is the saturation of phase j; ρ_j is the density of phase j; X_j^α is the mass fraction of component α in phase j; and F^α represents the advective mass flux of component α, expressed as

$$F^\alpha=\sum_j X_j^\alpha F_j=\sum_j X_j^\alpha\rho_j\boldsymbol{u}_j=\sum_j X_j^\alpha\left(-\boldsymbol{K}\frac{K_{rj}\rho_j}{\mu_j}\left(\nabla p_j-\rho_j\boldsymbol{g}\right)\right), \tag{2}$$

where F_j represents the flux of phase j; $\boldsymbol{u}_j$ is the Darcy velocity of phase j; $\boldsymbol{K}$ is the permeability tensor; K_{rj} is the relative permeability of phase j; μ_j is the viscosity of phase j; $\boldsymbol{g}$ is the gravitational acceleration; and p_j is the pressure of phase j.

Capillary pressure is taken into account in this system, a relationship between the gas phase pressure and the water phase pressure is expressed as follows:

$$p_g=p_w+P_c, \tag{3}$$

where P_c denotes the capillary pressure; p_g designates the gas phase pressure; p_w represents the water phase pressure.

For the coarse-scale system, the governing equations retain the same form as those in the fine-scale system, but the fine-scale parameters are replaced with corresponding upscaled parameters. We note that the fine- and coarse-scale two-phase governing equations are solved using the CO2STORE option in Eclipse.

2.2. Upscaling methodology

In this study, we develop both single-phase and two-phase upscaling procedures. The single-phase upscaling involves global flow-based upscaling methods for well index and permeability, while the two-phase upscaling includes global flow-based relative permeability upscaling and steady-state capillary pressure upscaling under the capillary limit (CL) assumption, which refers to a flow regime where capillary forces dominate over viscous forces in controlling fluid flow (Lohne and Virnovsky, 2006). The single-phase and two-phase upscaling procedures are introduced as

follows.

2.2.1. Single-phase upscaling

The single-phase upscaling procedure begins with a global single-phase fine-scale simulation. Then the upscaled well index can be calculated using the simulation results via:

$$WI_i^* = \frac{\sum_{l=1}^{N_w}\left(q_{well}^f\right)_l}{\langle p^f\rangle_i - \langle p_{well}^f\rangle_i}, \tag{4}$$

where N_w is the number of fine-scale well blocks within a single coarse block i; $\sum_{l=1}^{N_w}(q_{well}^f)_l$ is the total flow rate of all fine-scale wells within the coarse grid block i; $\langle p_{well}^f\rangle_i$ is the volume-averaged wellbore pressure; and $\langle p^f\rangle_i$ designates the volume-averaged pressure of fine grid blocks over the coarse well block i.

An example of the computation of WI^* is shown in Fig. 1. For simplicity, Cartesian uniform grids are employed. The dotted lines indicate the fine-scale grid blocks, while the solid lines represent the target coarse block. The circles denote the locations of fine-scale wells. In this example, nine fine-scale grid blocks are upscaled into a single coarse block, and the two fine-scale wells are merged into a single equivalent well within the coarse block. In this context, $\sum_{l=1}^{N_w}(q_{well}^f)_l$ is calculated as the sum of flow rates of the two fine-scale wells, $\langle p^f\rangle_i$ is determined as the average pressure of the nine fine-scale grid blocks, and $\langle p_{well}^f\rangle_i$ is computed as the average wellbore pressure of the two fine-scale wells.

Flow-based permeability upscaling is typically divided into the computation of upscaled permeability $\boldsymbol{K}^*$ or upscaled transmissibility T^*. Previous research (Chen et al., 2003; Romeu and Noetinger, 1995) has indicated that computing upscaled transmissibility T^* tends to be more accurate than computing upscaled permeability $\boldsymbol{K}^*$. Therefore, we also compute T^* using the following expression:

$$T_{i+\frac{1}{2}}^* = \frac{\langle q^f\rangle_{i+\frac{1}{2}}}{\langle p^f\rangle_i - \langle p^f\rangle_{i+1} - \rho g \Delta D^c}, \tag{5}$$

where i and $i+1$ refer to the indices of two adjacent coarse blocks, while $i+\frac{1}{2}$ indicates the index of the interface between these coarse blocks; $\langle q^f\rangle_{i+\frac{1}{2}}$ denotes the total flow rate across the fine grids at the coarse interface; $\langle p^f\rangle$ represents the volume-averaged fine-scale pressure over the coarse block; ρ is the density; g is the gravitational acceleration; and ΔD^c is the depth difference between coarse blocks i and $i+1$ (for the x and y directions, ΔD^c equals zero).

Fig. 2 illustrates an example of computing T^* in a two-dimensional (x–y) Cartesian coordinate system with uniform grids. The upscaled transmissibility T^* is anisotropic and must be computed separately for each direction (i.e., T_x^* and T_y^* are calculated independently). Taking T_y^* as an example (as shown in the right subplot), the total flow rate across the coarse interface is computed as $\langle q^f\rangle_{i+\frac{1}{2}} = \sum_{i=1}^{5} q_i$. The volume-averaged pressures $\langle p^f\rangle_i$ and $\langle p^f\rangle_{i+1}$ are calculated as the average pressures of the fine-scale grid blocks within the coarse blocks i and $i+1$, respectively.

Additionally, unphysical upscaled parameters (i.e., infinite values, not-a-numbers (NaNs), abnormal zeros, or negative values) may appear during flow-based upscaling computations (Chen, 2005; Li and Durlofsky, 2016), which usually cannot be accepted by numerical simulators. To address this issue, we develop replacement methods for the unphysical data.

For the unphysical upscaled well index, the equivalent permeability of the coarse well block is computed using the geometric mean method (Selvadurai and Selvadurai, 2014). The equivalent permeability is then applied to calculate the equivalent well index $WI(\boldsymbol{K}^*)$, which replaces the unphysical WI^*. For example, in the case of a vertical well in a three-dimensional model, the equivalent well index is computed as follows:

$$WI = \frac{2\pi\sqrt{k_x k_y}\Delta z}{\ln\frac{r_o}{r_w}}, \tag{6}$$

where r_o is the equivalent radius, defined as

$$r_o = 0.28\frac{\sqrt{\sqrt{\frac{k_y}{k_x}}\Delta x^2 + \sqrt{\frac{k_x}{k_y}}\Delta y^2}}{\sqrt[4]{\frac{k_y}{k_x}} + \sqrt[4]{\frac{k_x}{k_y}}}, \tag{7}$$

where r_w represents the wellbore radius; Δx, Δy, and Δz are the grid block dimensions; the terms k_x and k_y denote the permeabilities in the x and y directions, respectively.

For the unphysical upscaled transmissibility, the equivalent permeabilities of the two coarse blocks on both sides of the target coarse interface are also calculated using the geometric mean method (Selvadurai and Selvadurai, 2014). These equivalent

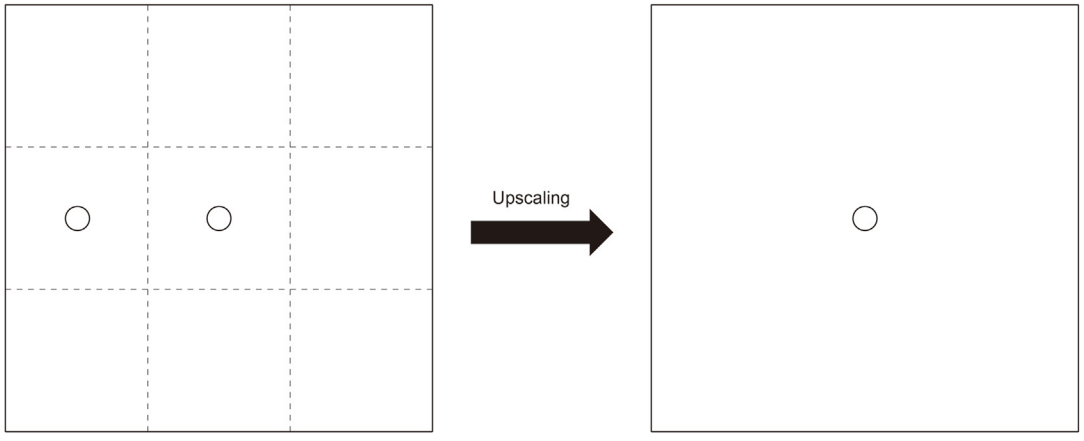


Fig. 1. Example of computing upscaled well index WI^*.

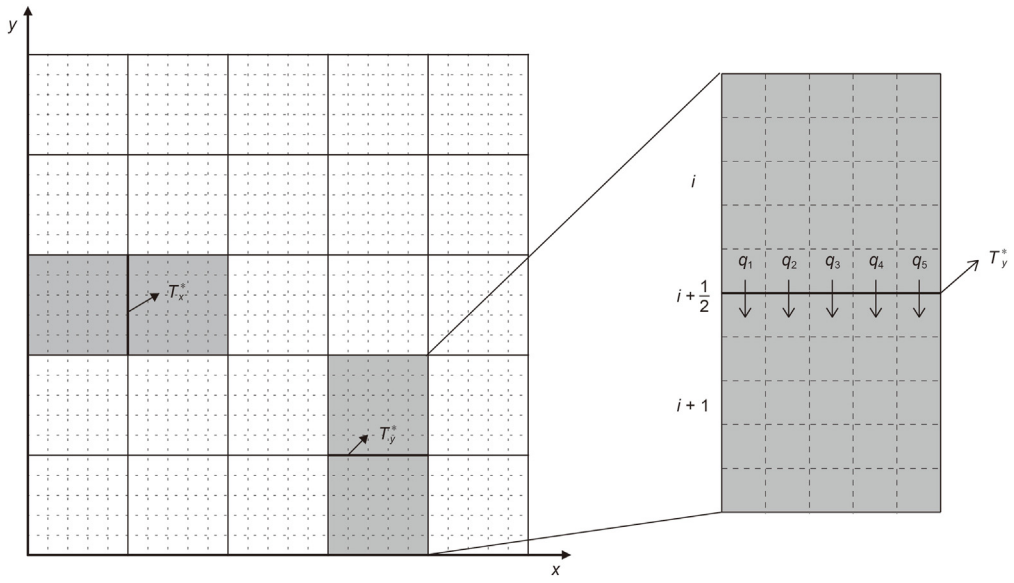


Fig. 2. Example of computing upscaled transmissibility T^*.

permeabilities are then employed to calculate the equivalent transmissibility, which replaces the unphysical T^*. For a scenario in the x direction, the calculations follow these equations:

$$k^{\mathrm{e}}_{i+\frac{1}{2}} = \frac{(\Delta x_i + \Delta x_{i+1}) k^{\mathrm{e}}_i k^{\mathrm{e}}_{i+1}}{\Delta x_{i+1} k^{\mathrm{e}}_i + \Delta x_i k^{\mathrm{e}}_{i+1}}, \tag{8}$$

$$T^{\mathrm{e}}_{i+\frac{1}{2}} = \frac{2 k^{\mathrm{e}}_{i+\frac{1}{2}} \Delta y \Delta z}{\Delta x_i + \Delta x_{i+1}}, \tag{9}$$

where k^{e}_i and k^{e}_{i+1} represent the equivalent permeabilities of coarse blocks i and $i+1$, respectively; $T^{\mathrm{e}}_{i+\frac{1}{2}}$ denotes the equivalent transmissibility at the target coarse interface; Δx, Δy, and Δz refer to the dimensions of the coarse block.

2.2.2. Two-phase upscaling

The global flow-based relative permeability upscaling initially entails a two-phase fine-scale simulation over the global domain. The results of this fine-scale simulation are then used to calculate the upscaled relative permeabilities through the following equations:

$$K^*_{\mathrm{rg}} = \frac{\langle \mu^{\mathrm{f}}_{\mathrm{g}} \rangle \langle q^{\mathrm{f}}_{\mathrm{g}} \rangle_{i+\frac{1}{2}}}{T^*_{i+\frac{1}{2}} \left(\langle p^{\mathrm{f}} \rangle_i - \langle p^{\mathrm{f}} \rangle_{i+1} - \rho^{\mathrm{c}}_{\mathrm{g}} g \Delta D^{\mathrm{c}} \right)}, \tag{10}$$

$$K^*_{\mathrm{rw}} = \frac{\langle \mu^{\mathrm{f}}_{\mathrm{w}} \rangle \langle q^{\mathrm{f}}_{\mathrm{w}} \rangle_{i+\frac{1}{2}}}{T^*_{i+\frac{1}{2}} \left(\left(\langle p^{\mathrm{f}} \rangle_i - P^*_{\mathrm{c},i} \right) - \left(\langle p^{\mathrm{f}} \rangle_{i+1} - P^*_{\mathrm{c},i+1} \right) - \rho^{\mathrm{c}}_{\mathrm{w}} g \Delta D^{\mathrm{c}} \right)}, \tag{11}$$

where i and $i+1$ represent two neighboring coarse blocks; $i+\frac{1}{2}$ indicates the coarse interface separating these coarse blocks; $T^*_{i+\frac{1}{2}}$ is the upscaled transmissibility defined at the coarse interface. The coarse-scale viscosities of the gas and water phases, represented by $\langle \mu^{\mathrm{f}}_{\mathrm{g}} \rangle$ and $\langle \mu^{\mathrm{f}}_{\mathrm{w}} \rangle$, are calculated as volume averages over the upstream coarse block. The volume-averaged pressure of fine gird blocks within coarse block i is denoted by $\langle p^{\mathrm{f}} \rangle_i$. $P^*_{\mathrm{c},i}$ represents the upscaled capillary pressure of coarse block i. $\langle q^{\mathrm{f}}_{\mathrm{g}} \rangle_{i+\frac{1}{2}}$ and $\langle q^{\mathrm{f}}_{\mathrm{w}} \rangle_{i+\frac{1}{2}}$ denote the sum of gas and water phase flow rates across the fine grids at the coarse interface, respectively. $\rho^{\mathrm{c}}_{\mathrm{g}}$ and $\rho^{\mathrm{c}}_{\mathrm{w}}$ represent the volume-averaged densities of the gas and water phases across the two coarse blocks that share the target interface. g denotes gravitational acceleration, and ΔD^{c} is the depth difference between coarse blocks i and $i+1$ (for the x and y directions, ΔD^{c} equals zero).

Similarly, the upscaled relative permeabilities K^*_{rg} and K^*_{rw} should be calculated anisotropically (as shown in Fig. 3). We note that the calculated K^*_{rg} and K^*_{rw} are defined at the coarse interface. However, since our simulator only accepts relative permeabilities defined on grid blocks, the computed K^*_{rg} and K^*_{rw} at the coarse interface are assigned to the upstream coarse block.

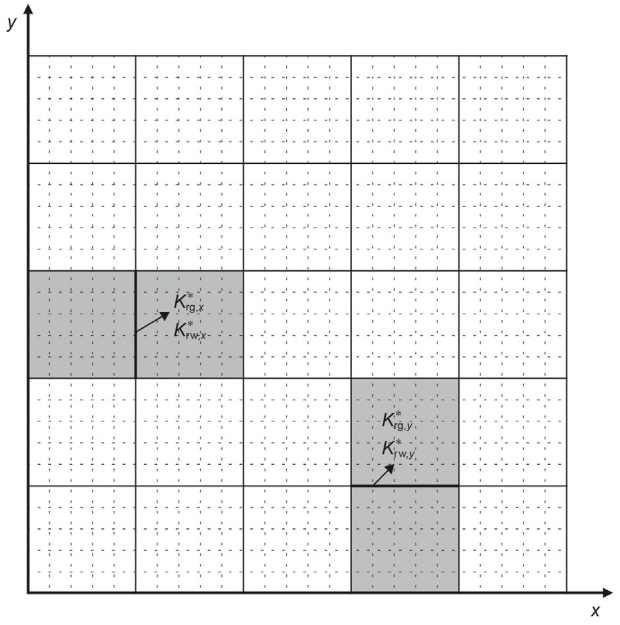


Fig. 3. Example of computing relative permeabilities K^*_{rg} and K^*_{rw}.

In our study, we discover that the presence of solubility greatly impacts the computation of upscaled relative permeability, leading to substantial errors. To mitigate this issue, simplified models that exclude solubility effects are used to perform the relative permeability upscaling computations. The resulting upscaled relative permeabilities are then applied to the coarse-scale full-physics model for numerical simulations. Similarly, flow-based relative permeability upscaling may generate unphysical data. We replace these unphysical data with fine-scale relative permeability curves.

The CL-based capillary pressure upscaling method demonstrates both high accuracy and computational efficiency among various existing capillary pressure upscaling methods (Cheng and Rabinovich, 2020). It works by inverting the saturation-capillary pressure relationship and averaging fine-scale saturation over the target coarse block under a range of capillary pressures. For detailed descriptions of this method, please refer to Rabinovich et al. (2016).

3. Geological models

The full-physics models are developed using the CO2STORE option in Eclipse (e300) to simulate CO_2 injection into saline aquifers. These models employ the CO_2-water equation of state provided by Spycher et al. (2003) and Spycher and Pruess (2005), which is applicable for temperatures within the range of 12 to 250 °C and pressures up to 600 bar. Our full-physics models account for relative permeability hysteresis, gravity, compressibility of rock and fluids, mutual solubility of CO_2 and brine, salt precipitation, and capillary heterogeneity. However, molecular diffusion and mineral trapping are excluded from the simulations, as their effects are considered negligible during the injection phase.

We consider two-dimensional models defined on an x–z Cartesian coordinate system. These models consist of 200× 200 grid blocks, with each grid block uniformly dimensioned at $\Delta x = \Delta y = \Delta z = 5$ m. The upscaling ratio is set to $10 \times 10 = 100$, resulting in coarse models composed of 20×10 coarse blocks. The reservoir is situated at a depth of 1000 m, leading to an initial pressure of 100 bar at the top boundary of the models. The reservoir temperature is kept constant at 55 °C. Under these conditions, CO_2 remains in a supercritical state. The rock compressibility is set to 4.934×10^{-5} bar^{-1}. The porosity is uniformly set to 0.25.

The permeability field is generated using Stanford Geostatistical Modeling Software (SGeMS) (Remy et al., 2009). We consider various types of permeability distributions, including Gaussian distributions with a range of correlation lengths, generated by the sequential Gaussian simulation (SGSIM) algorithm (Deutsch and Journel, 1992), and channelized distributions with horizontal and inclined channels, generated by the single normal equation simulation (SNESIM) algorithm (Strebelle, 2000, 2002). Samples of these models will be presented in Section 5.1. It should be noted that we assume an isotropic permeability field (i.e., $k_x = k_y$) in this study for simplicity.

The fluid components include water, CO_2, and NaCl. Water and CO_2 can exist in both liquid and gas phases, while NaCl can exist in both liquid and solid phases. The relative permeability curves are identical to those used in Zou and Durlofsky (2023), as shown in Fig. 4. Heterogeneous capillary pressure is defined using the Leverett-J function (Leverett, 1941):

$$P_c(S_g, k, \phi) = J(S_g)\sqrt{\frac{k_{ref}/\phi_{ref}}{k/\phi}}, \tag{12}$$

where k and ϕ denote the permeability and porosity; k_{ref} and ϕ_{ref} refer to the reference permeability and porosity; S_g represents the gas saturation; and $J(S_g)$ is the Leverett-J function. The capillary pressure curve calculated using a porosity of 0.25 and a permeability of 50 mD is presented in Fig. 5.

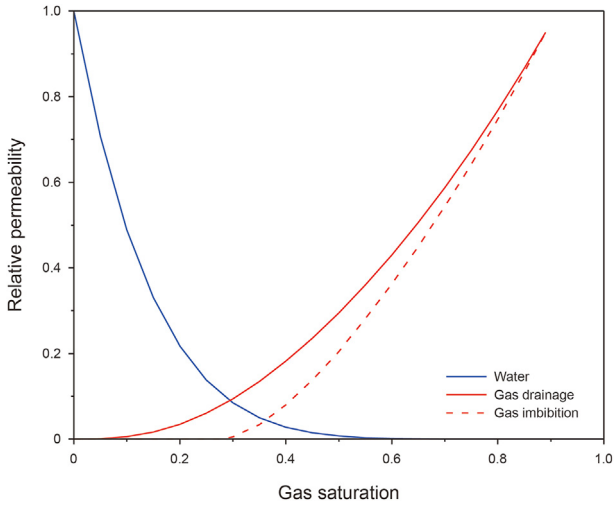


Fig. 4. Fine-scale relative permeability curves.

The models are initially saturated with brine containing 10,000 ppm NaCl and are driven by an injector with a constant CO_2 injection rate of 0.02 Mt/year and a producer with a constant bottomhole pressure (BHP) of 100 bar. We consider two basic flow patterns: left-to-right and bottom-to-top. In the left-to-right flow scenario (Fig. 6(a)), the CO_2 injector is positioned on the leftmost side of the model, while the producer is located on the rightmost side of the model, with both wells perforating the entire reservoir. In the bottom-to-top flow scenario (Fig. 6(b)), a horizontal injector is placed in the bottommost grid layer, while a horizontal producer is located in the topmost grid layer. Both wells are perforated along the full heel-to-toe trajectory. Samples of these two well trajectories are shown in Fig. 6.

Note that these geological models represent pilot tests rather than full-field CO_2 storage projects. Similar-sized models and well configurations have been used in previous CO_2 storage studies (Cameron and Durlofsky, 2012; Cheng and Rabinovich, 2020; Ide et al., 2007; Itthisawatpan, 2013; Juanes and MacMinn, 2008; Kou et al., 2022; Rabinovich et al., 2015). Additionally, this study serves as a preliminary investigation for deep learning-based

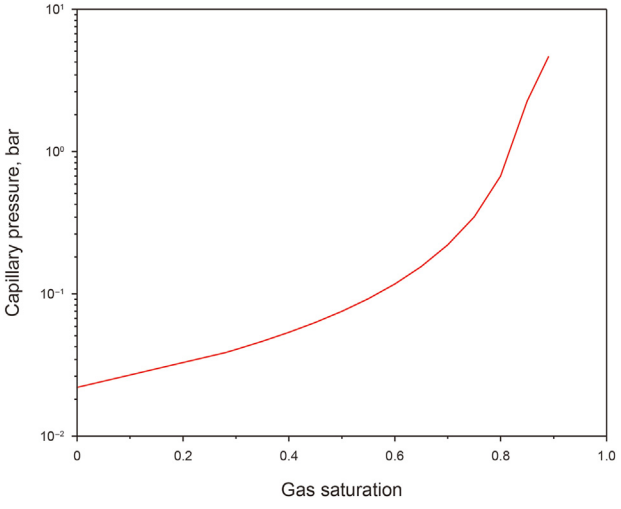
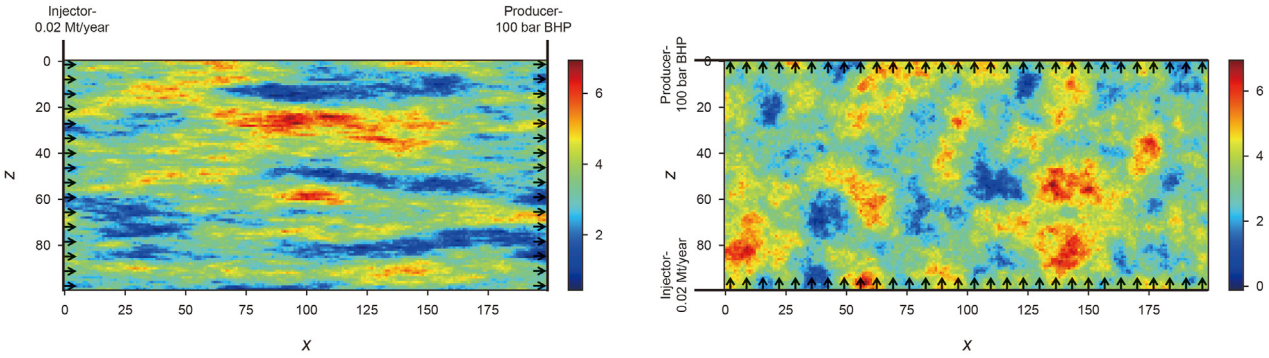


Fig. 5. Capillary pressure curve at porosity of 0.25 and permeability of 50 mD.



(a) Left-to-right flow **(b)** Bottom-to-top flow

Fig. 6. Schematics showing well configurations for left-to-right flow and bottom-to-top flow scenarios.

upscaling, with many parameters of the geological models fixed to reduce complexity. In future revisions, a broader range of parameters will be considered.

4. Deep learning surrogate models

In this section, we introduce the architectures of the surrogate models designed to replace the upscaling computations for well index, transmissibility, capillary pressure curve, and relative permeability curve. Due to the significant differences among these parameters, constructing and training a single surrogate model would be inefficient. Therefore, we develop dedicated surrogate models for each upscaled parameter. This allows us to employ simpler surrogate models that require less training time. Detailed descriptions are provided as follows.

4.1. Surrogate model for upscaled well index

The surrogate model for upscaled well index takes coarse well block properties as inputs and provides the predicted upscaled well index as output. The predictions can be expressed by

$$\widehat{y}_{WI^*} = f_{WI^*}(K_{\text{local}}, X_{\text{well}}, WI; \theta_{WI^*}), \tag{13}$$

where $\widehat{y}_{WI^*} \in \mathbb{R}$ represents the predicted upscaled well index; $K_{\text{local}} \in \mathbb{R}^{n_x \times n_z}$ denotes the local permeability distribution of the coarse well block, with n_x and n_z representing the dimensions of the local permeability in the x and z directions, respectively; $X_{\text{well}} \in \mathbb{R}$ represents the location of the coarse well block; $WI \in \mathbb{R}^{n_w}$ denotes the fine-scale well indices within the coarse well block with n_w referring to the number of fine-scale perforations; θ_{WI^*} represents the learnable parameters.

The upscaled well index surrogate model is constructed by integrating CNN and Transformer encoder. CNN is effective in capturing local spatial features through convolutional layers, making it particularly suited for processing grid-based inputs like permeability matrices. For more CNN applications in reservoir engineering, refer to Wen et al. (2021), Kwon et al. (2021), and Madasu et al. (2019). On the other hand, the Transformer encoder excels at handling long-range dependencies and contextual relationships within the data due to its self-attention mechanism (Wolf et al., 2020). By combining these two architectures, the model benefits from CNN's feature extraction capacity and the Transformer's global attention mechanism, resulting in more robust and accurate predictions.

The detailed architecture of this surrogate model is shown in Fig. 7. The model begins with a local permeability input (with a size of 10 × 10 in this study), which passes through two convolutional layers (Conv_1 and Conv_2), each followed by leaky ReLU activation and max pooling. The feature maps are progressively downsampled from a size of 32@10 × 10 to 64@2 × 2, after which they are flattened and passed through fully connected layers. Additional inputs, i.e., location of the coarse well block and fine-scale WI (with a size of 10 in this study), are also processed via fully connected layers. All inputs are then merged and fed into a Transformer encoder to produce the final upscaled well index (WI^*) output. The architecture of the Transformer encoder is shown in Fig. 8. The Transformer encoder consists of 3 layers, each equipped with a multi-head attention mechanism with 4 attention heads (nhead = 4). The feed forward neural network in each layer has a hidden dimension of 256 (dim = 256).

4.2. Surrogate model for upscaled transmissibility

The surrogate model for upscaled transmissibility accepts coarse block properties as inputs and produces the upscaled transmissibility predictions. The prediction process (i.e., f_{T^*}) can be expressed as follows:

$$\widehat{y}_{T^*} = f_{T^*}(K_{\text{local}}, X_{\text{inter}}, WI^*; \theta_{T^*}), \tag{14}$$

where $\widehat{y}_{T^*} \in \mathbb{R}$ represents the predicted upscaled transmissibility; $K_{\text{local}} \in \mathbb{R}^{n_x \times n_z}$ refers to the local permeability field of the target region, which includes two coarse blocks that share the coarse interface; n_x and n_z represent the number of fine grid blocks in the x and z directions within the target region; $X_{\text{inter}} \in \mathbb{R}$ represents the location of the coarse interface where T^* is defined; $WI^* \in \mathbb{R}$ represents the upscaled well index of the upstream coarse block at the coarse interface (if the upstream coarse block does not contain a well perforation, WI^* is set to 0); θ_{T^*} represents the trainable parameters.

Transmissibility upscaling shares similarities with well index upscaling, as both rely on global single-phase simulation. Consequently, we designed a surrogate model similar to the upscaled well index surrogate model, which integrates CNN and Transformer encoder. Taking the x-directional upscaled transmissibility in T_x^* as an example, Fig. 9 shows the architecture of the surrogate model. The Transformer encoder used in this model is identical to that shown in Fig. 8. The local permeability includes the two coarse blocks that share the coarse interface at which T^* is defined, thus the size of the local permeability in x direction is 20 × 10. For the z direction, the local permeability size is 10 × 20, and the surrogate model architecture remains unchanged. However, the dimensions of the local permeability input and the feature maps output by CNN

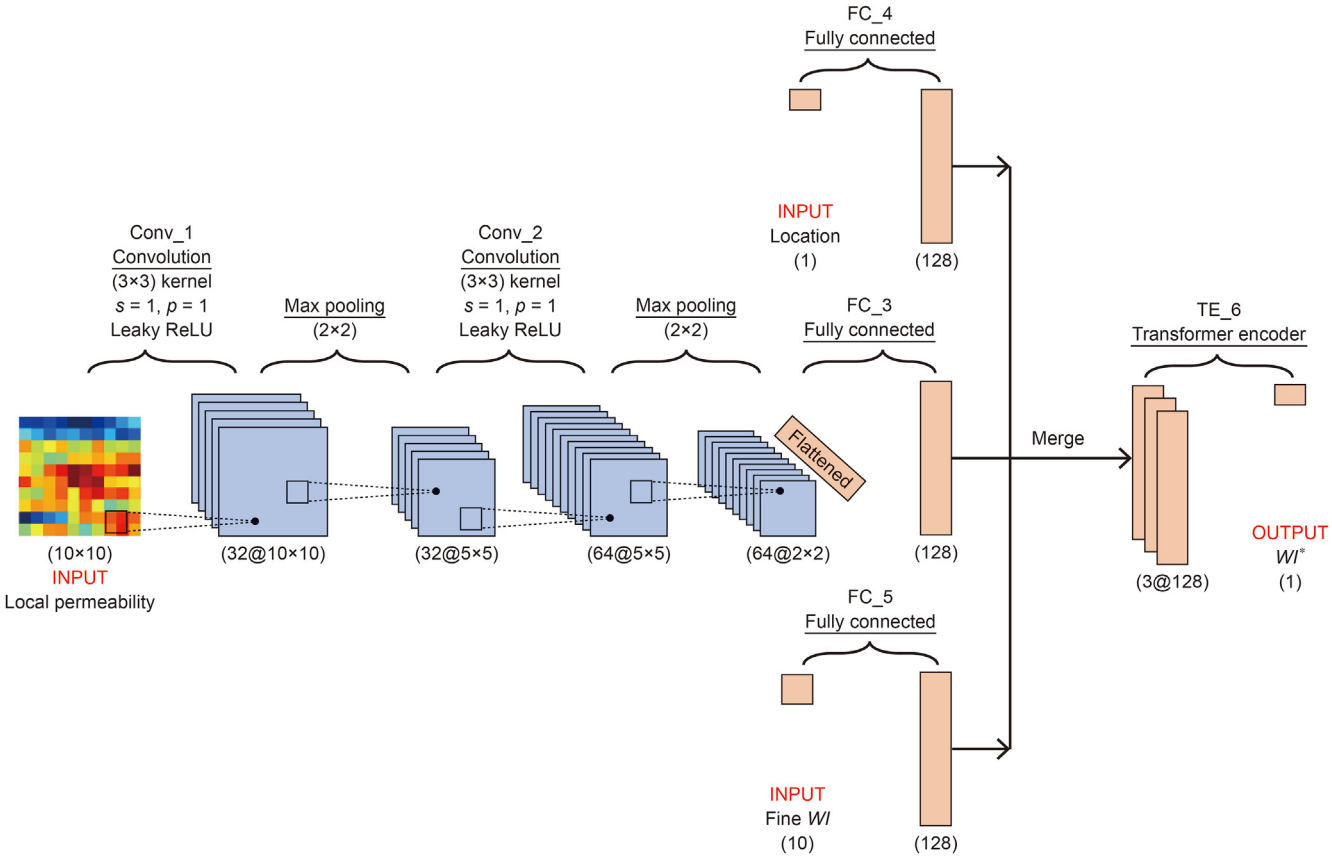


Fig. 7. Schematics showing the architecture of the surrogate model for upscaled well index.

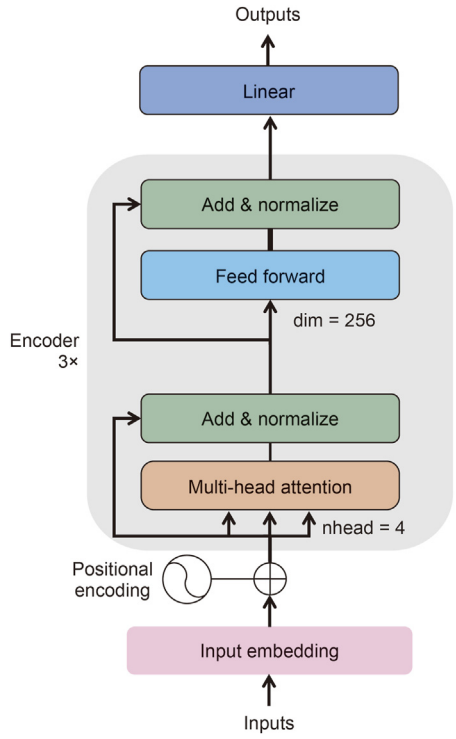


Fig. 8. Schematics showing the architecture of the Transformer encoder.

layers are swapped in height and width accordingly. Note that we train two separate surrogate models for the *x*- and *z*-directional upscaled transmissibility because the difference in input sizes increases the training complexity.

4.3. Surrogate model for upscaled capillary pressure

Fine-scale capillary pressure is defined using the Leverett-J function (Eq. (12)), which depends on permeability and porosity in this paper. Since we use a constant porosity, the surrogate model for upscaled capillary pressure requires only local permeability as input to predict the upscaled capillary pressure curve. To make it easier for the surrogate model to recognize capillary pressure-saturation curves, we discretized the curves into 12 points, with each point corresponding to a specific saturation value. The prediction process, denoted as $f_{P_c^*}$, can be expressed as

$$\hat{y}_{P_c^*} = f_{P_c^*}\left(K_{\text{local}}; \theta_{P_c^*}\right), \tag{15}$$

where $\hat{y}_{P_c^*} \in \mathbb{R}^{12}$ represents the predicted upscaled capillary pressure points; $K_{\text{local}} \in \mathbb{R}^{n_x \times n_z}$ denotes the local permeability of the target coarse block, with n_x and n_z representing the block's dimensions in the *x* and *z* directions, respectively; and $\theta_{P_c^*}$ represents the trainable parameters.

Fig. 10 shows the architecture of the surrogate model for upscaled capillary pressure. The model takes local permeability (of size 10×10) as input and processes it through CNN layers followed by fully connected layers. The CNN extracts relevant features from the permeability data, while the fully connected layers output the

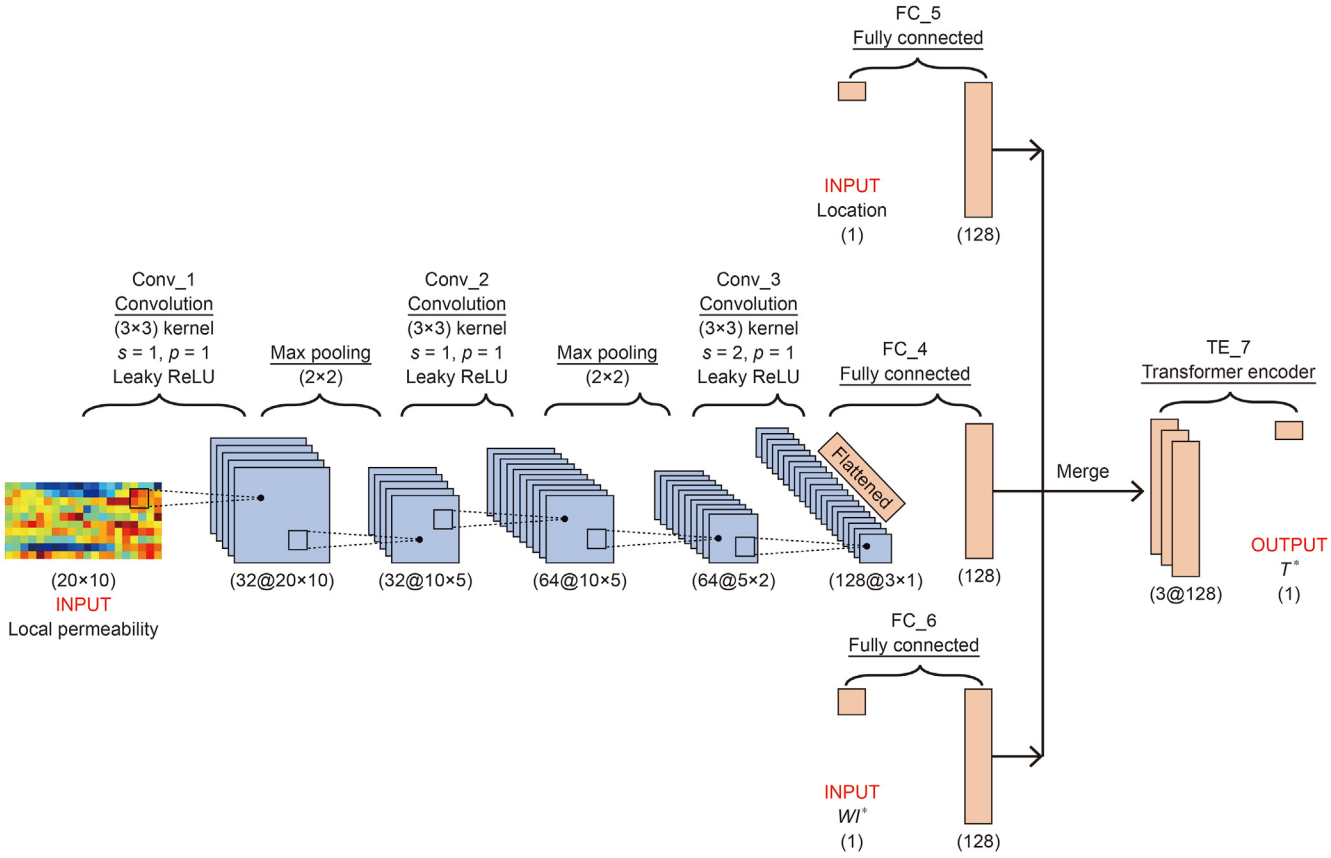


Fig. 9. Schematics showing the architecture of the surrogate model for upscaled transmissibility.

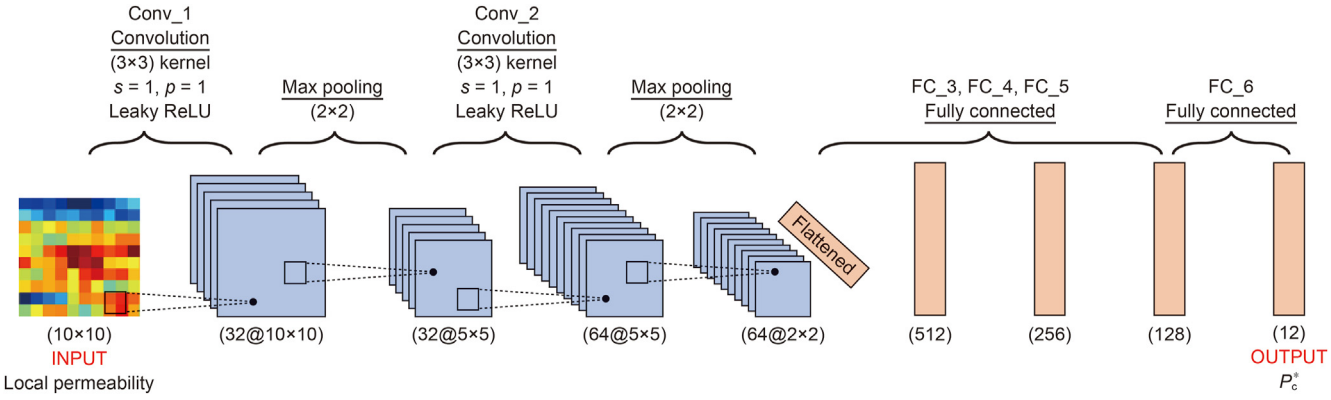


Fig. 10. Schematics showing the architecture of the surrogate model for upscaled capillary pressure.

discretized capillary pressure values, which are then combined to form the complete capillary pressure curve.

4.4. Surrogate model for upscaled relative permeability

The surrogate model for upscaled relative permeability takes the properties of the target coarse block and the upscaled parameters predicted by other surrogate models described earlier as input and predicts the upscaled relative permeability. Similar to the approach used for capillary pressure, we discretized the relative permeability-saturation curves into 12 points, each corresponding to a specific saturation value. The prediction can be expressed as follows:

$$\hat{y}_{K_r^*} = f_{K_r^*}\left(K_{\text{local}}, X_{\text{inter}}, WI^*, T^*, P_c^*; \theta_{K_r^*}\right), \tag{16}$$

where $\hat{y}_{K_r^*} \in \mathbb{R}^{12}$ represents the predicted upscaled relative permeability points; $K_{\text{local}} \in \mathbb{R}^{n_x \times n_z}$ represents the local permeability, which includes two coarse blocks sharing the target coarse interface. Similarly, n_x and n_z represent the dimensions of the local permeability. Specifically, for the x direction, $n_x = 20$ and $n_z = 10$, while for the z direction, $n_x = 10$ and $n_z = 20$. $X_{\text{inter}} \in \mathbb{R}$ denotes the location of the target coarse interface. $WI^* \in \mathbb{R}$ represents the upscaled well index of the upstream coarse block at the coarse interface. If the upstream coarse block lacks a well perforation, WI^*

is assigned a value of 0. $T^* \in \mathbb{R}$ denotes the upscaled transmissibility, and $P_c^* \in \mathbb{R}^{12}$ denotes the predicted upscaled capillary pressure points. $\theta_{K_r^*}$ represents the learnable parameters.

Relative permeability upscaling is the most complex part of the entire upscaling procedure, as it involves solving partial differential equations (PDEs) for intricate two-phase flow problems. Furthermore, relative permeability is a dynamic property that changes with saturation (or time), which adds additional difficulty for deep learning models to capture this continuous variation. To improve accuracy, we developed a surrogate model that integrates CNN, Transformer encoder, and FNO. FNO provides significant advantages in addressing PDE-related problems. It uses Fourier transforms to map spatiotemporal data into the frequency domain, making it particularly well-suited for handling the dynamic changes and complex interactions in two-phase flow problems. For more discussion about FNO, please refer to Li et al. (2020, 2023).

Fig. 11 presents the architecture of the surrogate model for upscaled relative permeability (again taking the x direction as an example). The Transformer encoder architecture is identical to that shown in Fig. 8, while the FNO architecture is illustrated in Fig. 12. To reduce training complexity, the relative permeabilities for the gas phase and water phase are trained separately in the x and z directions. For the z direction, the height and width of the local permeability and the outputs of CNN layers are swapped. As a result, there are a total of four surrogate models for upscaled relative permeabilities (i.e., $K^*_{\mathrm{rw},x}$; $K^*_{\mathrm{rw},z}$; $K^*_{\mathrm{rg},x}$; $K^*_{\mathrm{rg},z}$).

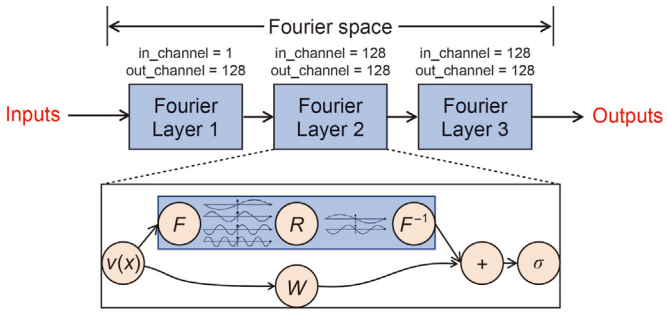


Fig. 12. Schematics showing the architecture of the FNO.

5. Overall workflow

5.1. Datasets generation

Eight types of permeability distributions are considered, including two types of channelized models with horizontal and inclined channels, and six types of Gaussian models with varying correlation lengths. For each type, 1000 models are randomly generated following the descriptions in Section 3, resulting in a total of 8000 different permeability distributions. Fig. 13 illustrates the examples of the eight types of permeability (log$k$) distributions, where l_x and l_z represent the correlation lengths in the x and z directions, respectively.

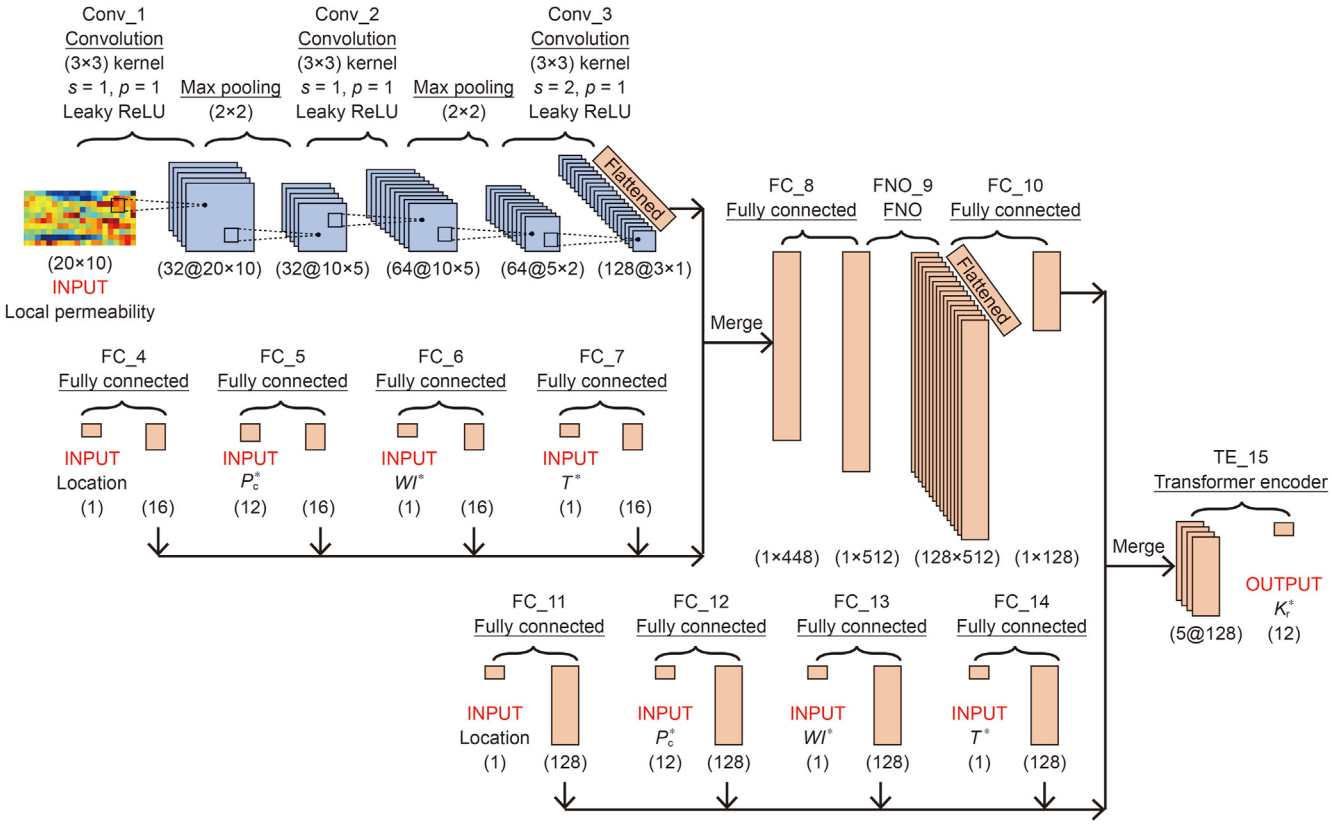


Fig. 11. Schematics showing the architecture of the surrogate model for upscaled relative permeability.

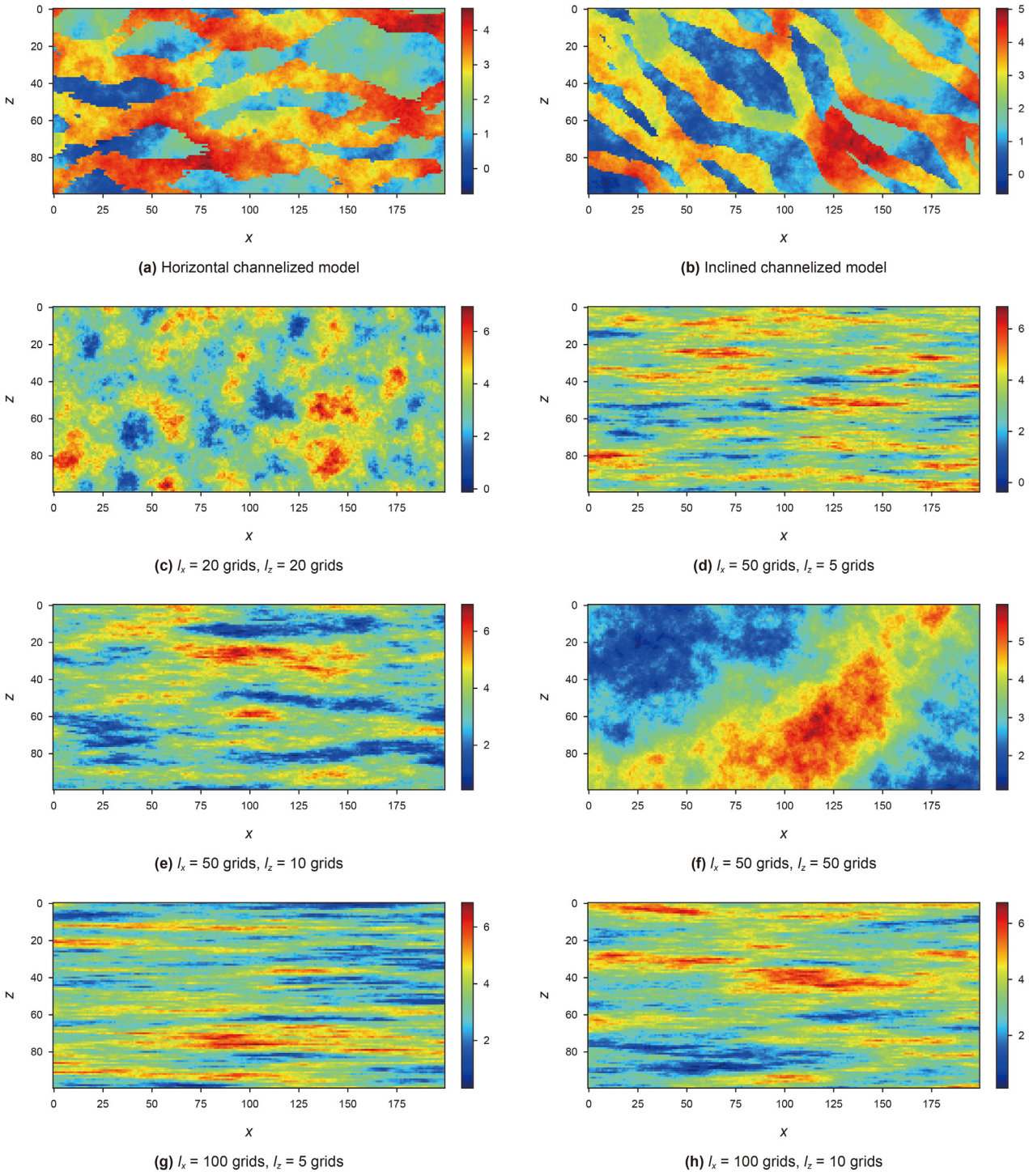


Fig. 13. Schematics showing examples of the eight types of permeability (log*k*) distributions.

5.2. Data preprocessing

Previous studies have shown that upscaled parameters calculated from coarse blocks (or interfaces) with low flow rates can be inaccurate (Chen, 2005; Chen et al., 2003; Durlofsky, 2005). Using these inaccurate upscaled parameters can degrade the quality of the datasets, leading to a decline in the performance of surrogate models. We categorize the inaccurate upscaled parameters into the following three types and apply different handling methods.

The first type is unphysical data, as mentioned earlier, which refers to upscaled parameters with obvious errors that cannot be accepted by numerical simulators, such as infinite values, NaNs, abnormal zeros, or negative values. Unphysical data are easy to identify and are typically replaced with computed alternative data. However, since these replacements do not accurately represent the upscaling results, all such replacement values should be excluded when constructing datasets for deep learning models.

The second type is outlier data, which refers to upscaled

parameters that differ significantly from typical values but do not exhibit the obvious errors mentioned above. Here, we employ the isolation forest algorithm (Liu et al., 2008) to filter the outliers.

The third type is noise data, which refers to upscaled parameters with certain errors but mixed with typical values, without showing outlier characteristics. Noise data are difficult to identify. We mitigate their impact by enhancing the generalization ability of the deep learning model, such as by increasing the number of training samples and strictly avoiding overfitting.

Additionally, we perform a logarithmic transformation $\ln(x+10^{-6})$ to both the input and output data. This transformation enhances the contrast in small values, allowing the surrogate model to more effectively capture and learn these variations, which improves training efficiency and model performance. During the inference process, the predictions need to be reverted to their original scale using the inverse logarithmic function.

5.3. Training and inference strategies

The 8000 geological models described in Section 5.1 are split into training, validation, and test sets with ratios of 40%, 10%, and 50%, respectively. Data preprocessing described in Section 5.2 are applied to both the training and validation sets. During the training process, the batch size is set to 256, and the learning rate is initialized to 10^{-4}. The mean square error (MSE) loss function is employed

$$L_{\mathrm{MSE}} = \frac{1}{N}\sum_{i=1}^{N} \|y_i - \widehat{y}_i\|_2^2, \tag{17}$$

where N is the number of training samples within a batch; y represents the true data; and $\widehat{y}$ is the prediction from the surrogate model. The Adam optimizer (Kingma, 2014) is used to minimize the loss function. The surrogate models are trained on an NVIDIA A100 GPU.

Fig. 14 illustrates the workflow of the inference process. The input data, including geological model and well configuration information, first pass through a feature extractor, which identifies and extracts relevant features for each surrogate model. These features are then passed into the trained surrogate models. Finally, the prediction outputs from these surrogate models are aggregated to construct coarse-scale models. Note that prediction values smaller than 10^{-6} are treated as zero.

Table 1
Performance metrics for surrogate models on training and validation sets.

Surrogate model	Error metric	Training set	Validation set
WI^*	R^2	0.9999	0.9999
	MSE	0.0021	0.0020
T_x^*	R^2	0.9996	0.9996
	MSE	0.0035	0.0036
T_z^*	R^2	0.9998	0.9998
	MSE	0.0039	0.0041
P_c^*	R^2	1.0000	1.0000
	MSE	0.0001	0.0001
$K_{rw,x}^*$	R^2	0.9900	0.9886
	MSE	0.1692	0.1928
$K_{rg,x}^*$	R^2	0.9803	0.9786
	MSE	0.3338	0.3512
$K_{rw,z}^*$	R^2	0.9772	0.9764
	MSE	0.4295	0.4423
$K_{rg,z}^*$	R^2	0.9791	0.9761
	MSE	0.3650	0.3713

6. Results

To investigate the effect of flow patterns on DL-based predictions, we consider two flow patterns: left-to-right (Fig. 6(a)) and bottom-to-top (Fig. 6(b)). These flow patterns are analyzed in the following two cases.

6.1. Case 1: left-to-right flow scenario

6.1.1. Performance evaluation of surrogate models

The first case involves the left-to-right flow pattern. We first evaluate the performance of the DL-based surrogate models. Table 1 presents the performance metrics for surrogate models on training and validation sets. The error metrics represent the average errors of all samples. The MSE is defined by Eq. (17), while R^2 score is given by

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{18}$$

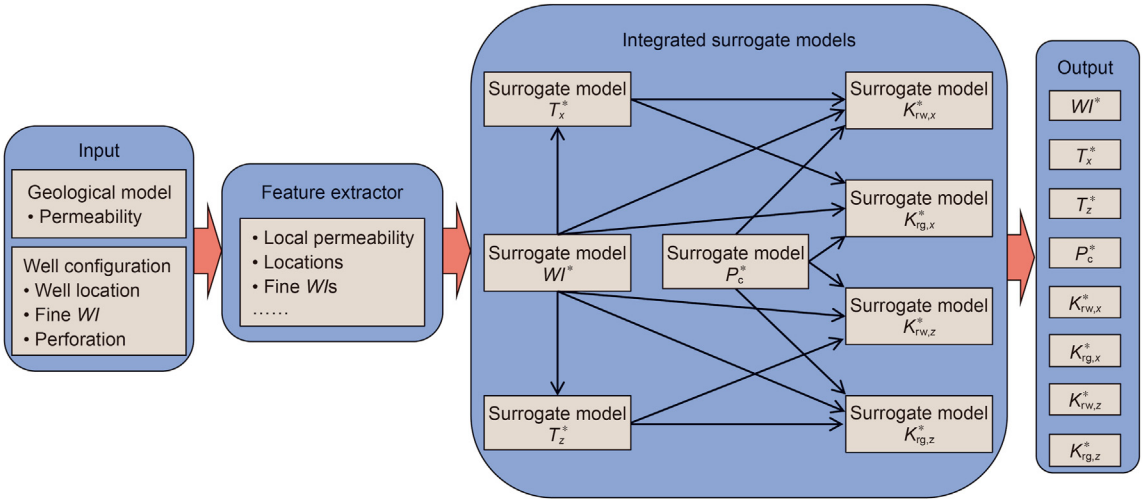


Fig. 14. Schematic showing the workflow of the inference process.

where n is the number of samples; y_i is the true value; $\hat{y}_i$ is the predicted value; and $\bar{y}$ is the mean of the true values.

The surrogate models for WI^*, T_x^*, T_z^*, and P_c^* demonstrate excellent accuracy, with R^2 scores close to 1 and very low MSE errors across both training and validation sets, indicating strong generalization capabilities. In contrast, the surrogate models related to upscaled relative permeabilities (i.e., $K_{rw,x}^*$; $K_{rg,x}^*$; $K_{rw,z}^*$; $K_{rg,z}^*$) exhibit slightly lower performance. Therefore, we primarily focus on analyzing the errors associated with the predicted upscaled relative permeabilities. Fig. 15 shows samples of upscaled relative permeabilities with P90 (90th percentile) MSE error in the validation set. The black solid curves represent the true curves computed by traditional flow-based upscaling, while the red dashed curves denote the predicted curves from DL-based surrogate models. The results show that the predicted curves closely match the true curves, particularly for those in the x direction. In contrast, the predicted curves in the z direction exhibit larger errors. We believe this is because, in this case, the fluid primarily flows in the x direction, while the flow in the z direction is relatively weak. As a result, a higher proportion of noise data is generated in the z direction, which affects the deep learning model's performance. However, since the z direction is not the main flow direction, the impact of these errors on the overall numerical simulation results is minimal. As the figures show P90 errors, 90% of the samples in the validation set have MSE errors lower than those shown in the figures. This demonstrates that the DL-based predictions exhibit very high accuracy and robustness.

6.1.2. Use of surrogate models for upscaling

We now apply the trained surrogate models to perform upscaling on the test set, which consists of 4000 geological models. First, the surrogate models are used to predict the upscaled parameters for the coarse-scale models, after which numerical simulations are performed for fine- and coarse-scale models. Subsequently, using the fine-scale simulation results as benchmarks, we respectively calculate the relative errors for the traditional upscaling models and the DL-based upscaling models. The relative errors in flow rates are calculated via

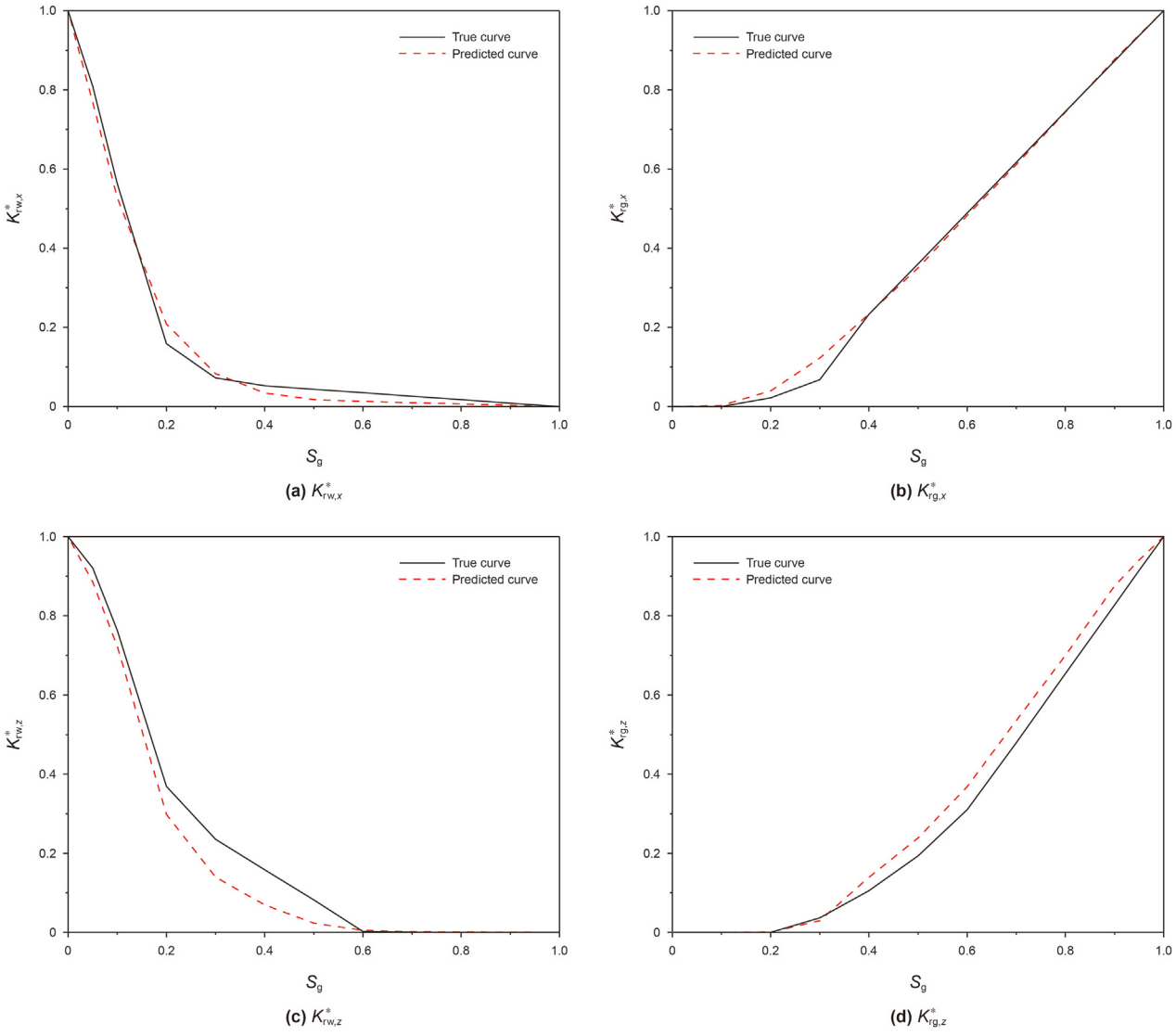


Fig. 15. Schematics showing $K_{rw,x}^*$, $K_{rg,x}^*$, $K_{rw,z}^*$ and $K_{rg,z}^*$ samples with P90 MSE error in the validation set.

$$e_1 = \frac{\int_{PVI} |y^{\text{fine}} - y^{\text{coarse}}| \mathrm{d}PVI}{\int_{PVI} y^{\text{fine}} \mathrm{d}PVI}, \tag{19}$$

while for pressure or saturation fields, the relative errors can be calculated by

$$e_2 = \frac{1}{N}\sum_{j=1}^{N} \frac{\int_{PVI} \left| \langle y_j^{\text{fine}} \rangle - y_j^{\text{coarse}} \right| \mathrm{d}PVI}{\int_{PVI} \langle y_j^{\text{fine}} \rangle \mathrm{d}PVI}, \tag{20}$$

where y^{fine} represents the fine-scale results; y^{coarse} denotes coarse-scale results; N is the number of coarse blocks in the global domain; $\langle y_j^{\text{fine}} \rangle$ represents the volume-averaged pressure or saturation of fine grid blocks within the target coarse block; y_j^{coarse} denotes the pressure or saturation in the target coarse block; and PVI represents pore volume injected.

Fig. 16 presents box plots showing the relative errors of gas (Q_g) and water (Q_w) phase flow rates at the producer for traditional and DL-based upscaling models. The red box plots represent the traditional upscaling model, while the blue box plots correspond to the DL-based upscaling model. The labels a–h on the horizontal axis correspond to eight permeability fields (examples shown in Fig. 13), and "total" represents the overall statistical error across all 4000 models in the test set. Each box plot summarizes the distribution of relative errors, where the entire box represents the range from the 10th percentile (P10) to the 90th percentile (P90) of the data. The positions of key percentiles (e.g., P25 and P75) are illustrated in the green legend. The solid line inside the box indicates the median relative error, the square represents the mean, and the whiskers extend to 1.5 times the interquartile range (IQR), which represents the difference between P25 and P75 of the data.

It is evident that DL-based upscaling model generally outperforms the traditional upscaling model in terms of reducing the relative errors of water phase flow rates across all permeability types (Fig. 16(a)). The DL-based upscaling model consistently shows narrower error distributions, indicating more stable performance with smaller variations. This suggests that the DL-based model delivers more reliable and accurate predictions, while the traditional model tends to have greater variability and higher potential for larger errors.

The box plots in Fig. 16(b) compare the relative errors of Q_w between traditional and DL-based upscaling models. The DL-based upscaling models generally exhibit narrower error ranges and whiskers, indicating lower variability in predictions. Furthermore, the DL-based upscaling shows a clear advantage at higher percentiles (i.e., P90 and P75), where it outperforms the traditional upscaling. However, at the lower percentiles (i.e., P25 and P10), no significant improvement is observed with the DL-based upscaling models compared to the traditional upscaling models. Overall, the DL-based upscaling models demonstrate superior performance, providing better accuracy and stability in most scenarios.

Figs. 17 and 18 present gas and water flow rate samples with relative errors of P90, P75, P50, and P25. The P90 percentile is critical for evaluating high-error scenarios in model performance, indicating that 90% of the models produce more accurate results than the sample at this level. The same interpretation applies to the other percentiles. Notably, even at P90, the models demonstrate reasonable accuracy, indicating that most models can achieve reliable results using DL-based surrogate models.

Fig. 19 presents box plots showing the relative errors of gas saturation and pressure fields for traditional and DL-based upscaling models. The description of box plots is consistent with the previous description. It is evident that the DL-based upscaling model (in blue) consistently exhibits lower relative errors and variability than the traditional upscaling model (in red).

We next analyze the field results generated by DL-based upscaling models. Figs. 20 and 21 illustrate gas saturation field samples with relative errors of P90 and P50, respectively. Each sample is shown for two injection stages: $PVI = 0.1$ (before gas breakthrough) and $PVI = 0.5$ (after gas breakthrough). The samples presented in these figures demonstrate reasonable accuracy in the gas saturation fields. Moreover, 90% of the models in the test set outperform the results shown in Figs. 20, and 50% of the models are more accurate than those shown in Fig. 21, indicating that most DL-based models can reliably provide accurate gas saturation field results.

The relative errors in the pressure field are lower than those of the gas saturation field, as shown by comparing error values in Fig. 19(a) and (b). Fig. 22 presents a pressure field sample with a relative error of P90. The sample presented in the figure exhibits very high accuracy, indicating the majority of DL-based models can achieve highly accurate pressure field results.

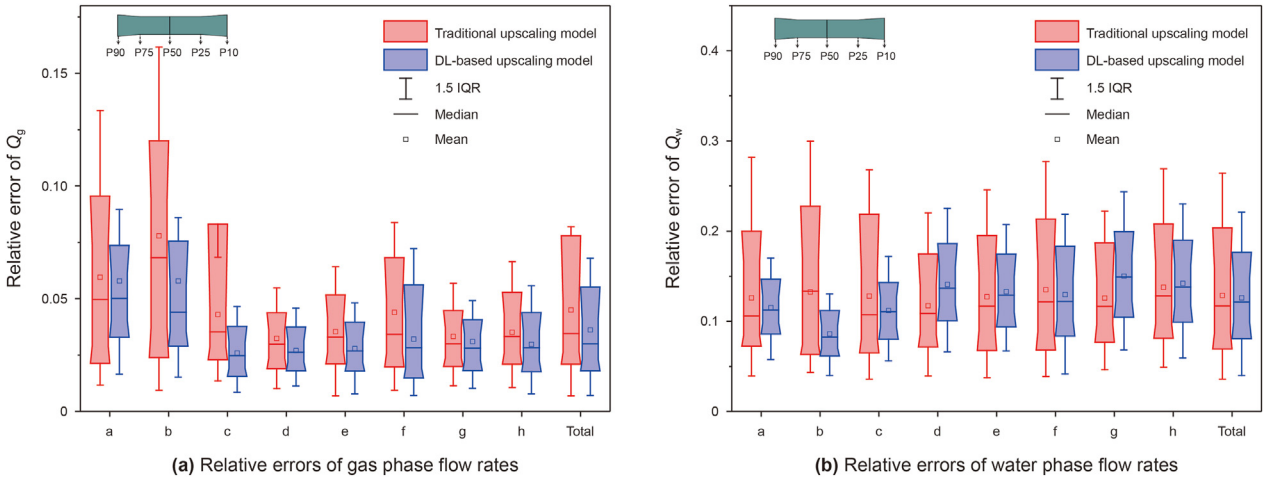


Fig. 16. Box plots showing the relative errors of gas phase and water phase flow rates at the producer for traditional and DL-based upscaling models.

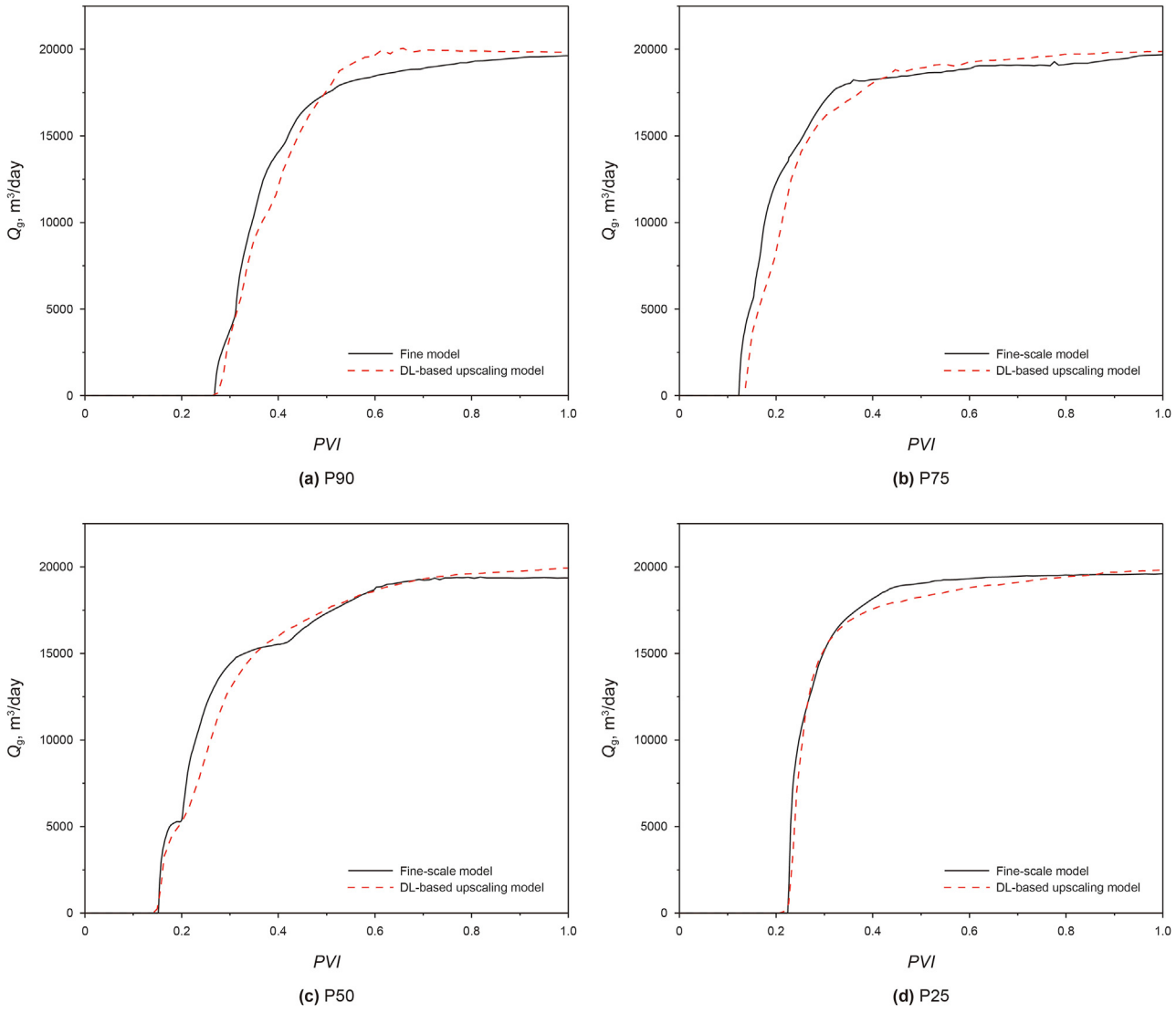


Fig. 17. Gas flow rate samples with relative errors of P90, P75, P50, and P25.

6.2. Case 2: bottom-to-top flow scenario

6.2.1. Performance evaluation of surrogate models

The second case involves a bottom-to-top flow scenario. We first evaluate the performance of the DL-based surrogate models in this flow scenario. Table 2 shows the performance summary for surrogate models on training and validation sets. The MSE error and R^2 score are calculated by Eqs. (17) and (18), respectively. The surrogate models for WI^*, T_x^*, T_z^*, and P_c^* again exhibit excellent accuracy, with R^2 scores exceeding 0.999 and MSE errors below 0.005 for both training and validation sets. Therefore, we primarily concentrate on surrogate models for upscaled relative permeabilities (i.e., $K_{rw,x}^*$; $K_{rg,x}^*$; $K_{rw,z}^*$; $K_{rg,z}^*$), which are comparatively less accurate than the aforementioned models. Fig. 23 illustrates samples of upscaled relative permeabilities with P90 MSE error in the validation set, which indicates that 90% of the upscaled relative permeability predictions are more accurate than the samples shown in the figure. By analyzing the data from Table 2 and Fig. 23, it is evident that the predictions in the z direction are highly accurate, while the accuracy in the x direction is comparatively lower. We attribute this to the weaker flow in the x direction, as the z direction is the primary flow direction. As discussed earlier, weaker flows tend to introduce a higher portion of noise data, which reduces the performance of the surrogate models. However, errors in the non-primary flow direction have a limited impact on the global simulation results. Overall, the surrogate models for upscaled relative permeabilities are able to provide reliable predictions.

6.2.2. Use of surrogate models for upscaling

We next apply the trained surrogate models to geological models in the test set. Similarly, the relative errors of coarse-scale simulation results are calculated using Eqs. (19) and (20). Fig. 24 illustrates box plots that illustrate the relative errors of gas and water phase flow rates at the producer for traditional and DL-based upscaling models. As described earlier, these box plots summarize the distribution of relative errors across a–h types of permeability fields, as well as the overall relative error. The red box plots represent the traditional upscaling model, while the blue box plots show the DL-based upscaling model. For the gas flow rates, the DL-based upscaling models generally exhibit lower and more stable relative errors compared to the traditional models. However, for the water flow rates, the DL-based upscaling models provide no significant improvement.

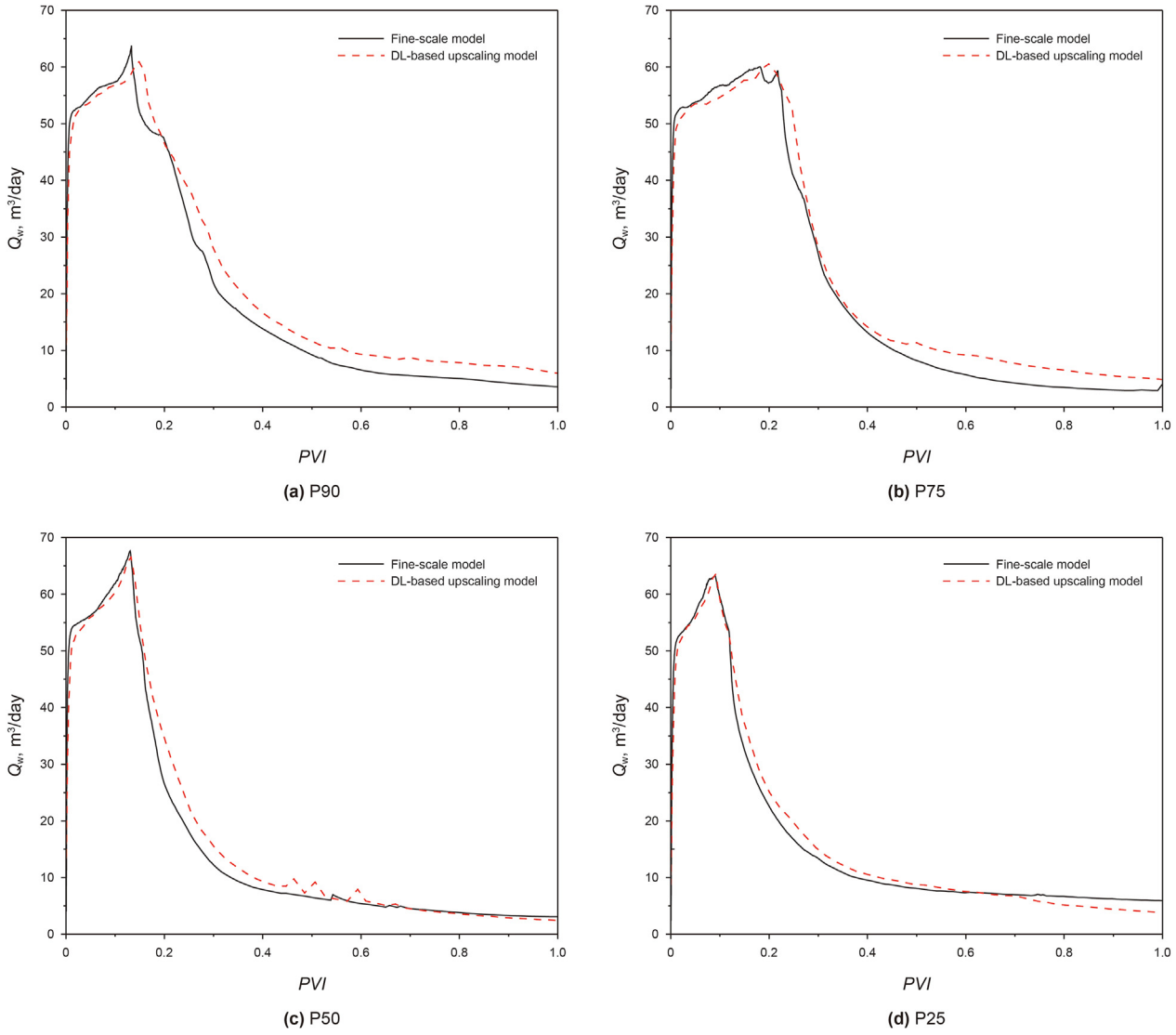


Fig. 18. Water flow rate samples with relative errors of P90, P75, P50, and P25.

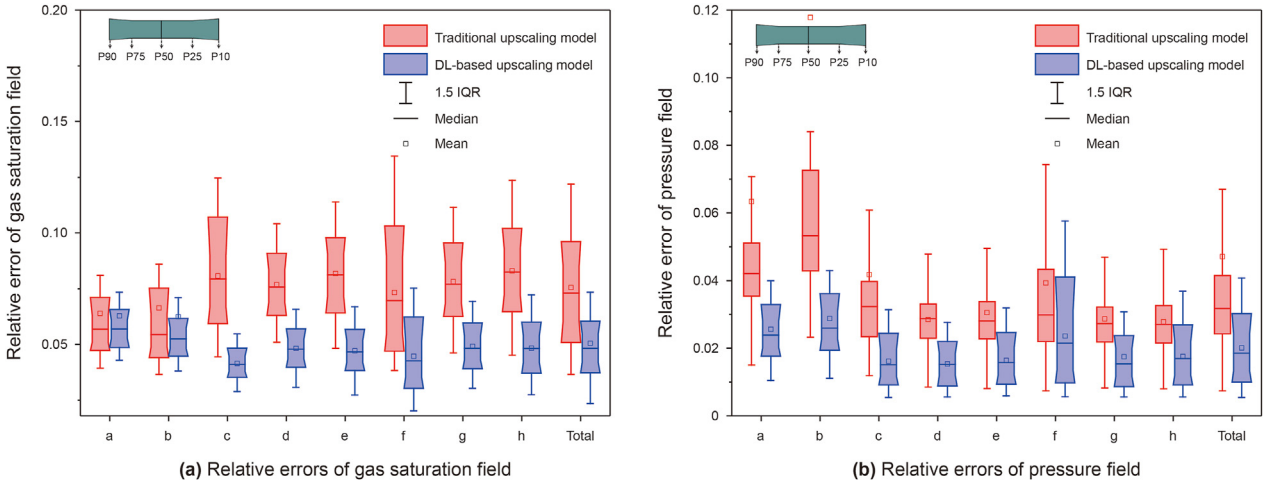
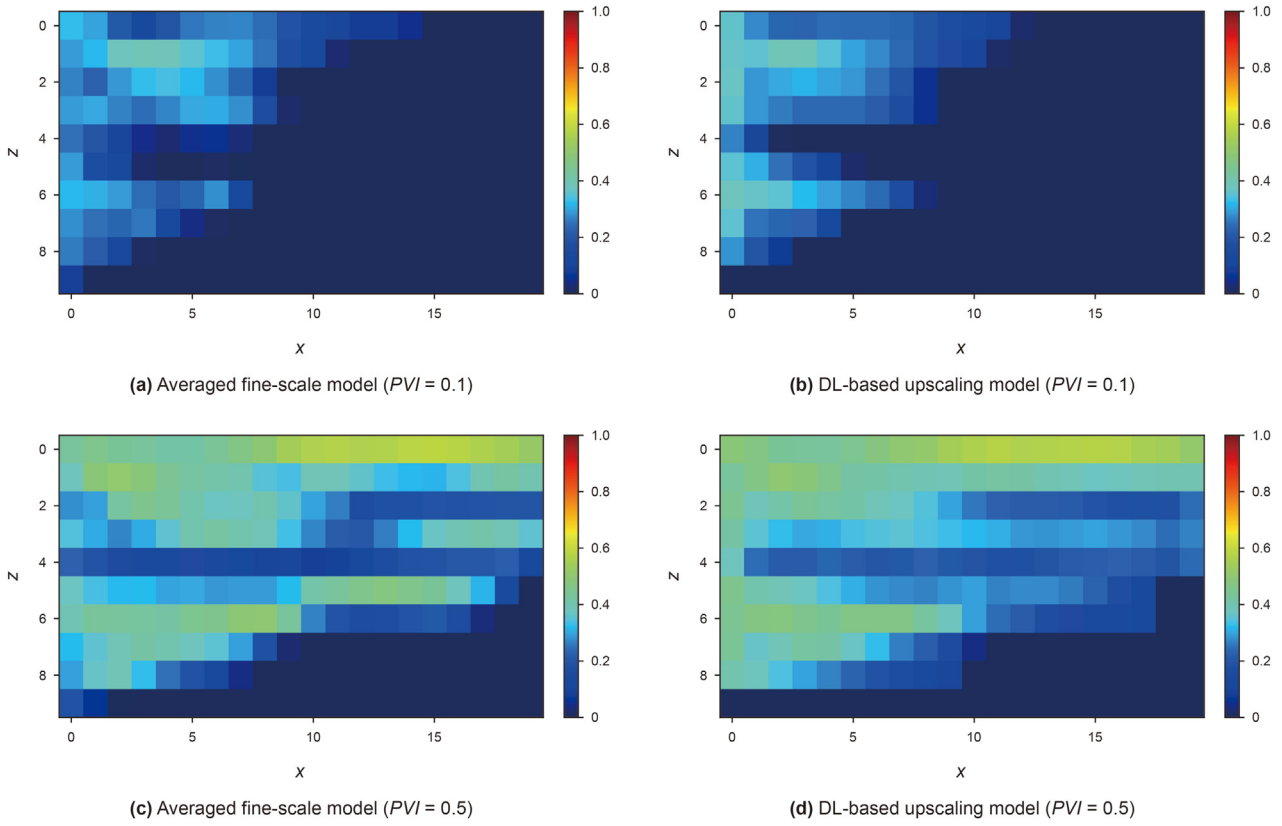


Fig. 19. Box plots showing the relative errors of gas saturation and pressure fields for traditional and DL-based upscaling models.



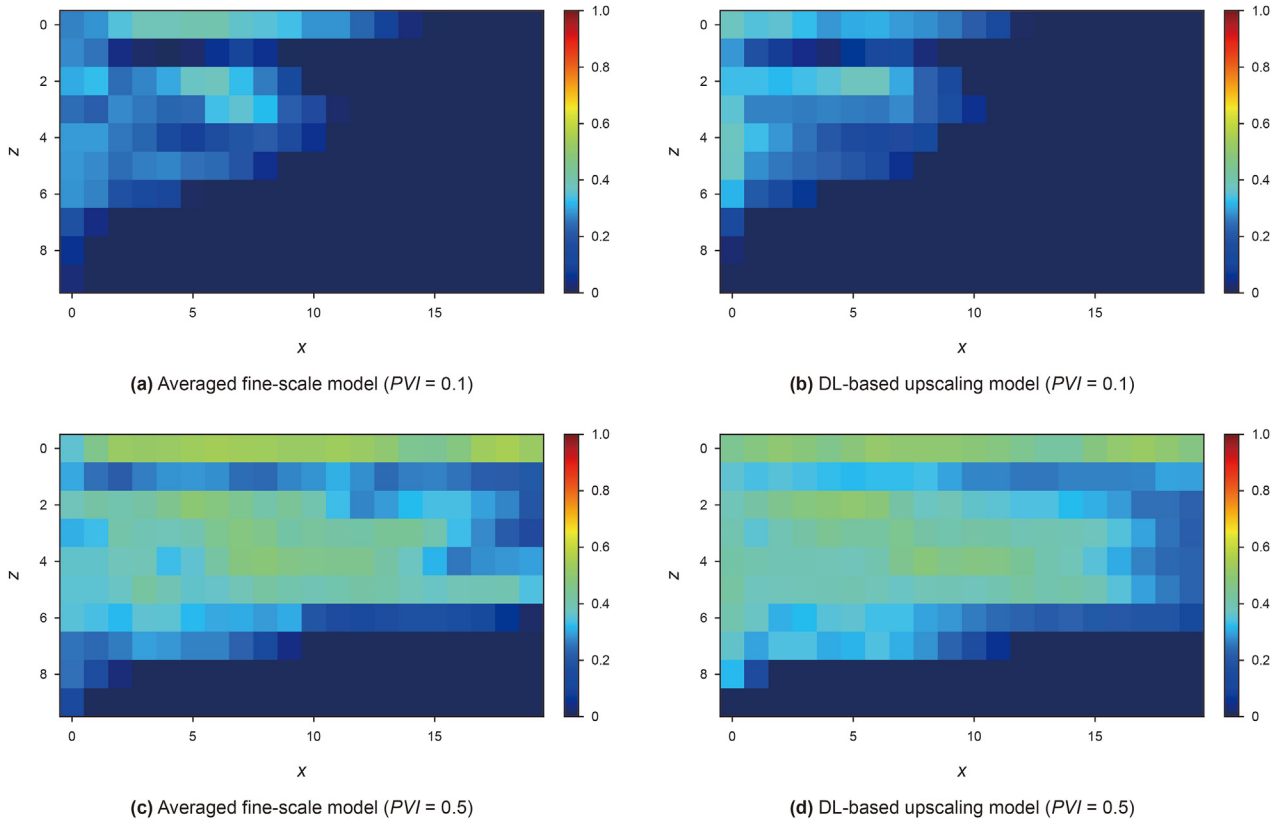
(a) Averaged fine-scale model (*PVI* = 0.1)

(b) DL-based upscaling model (*PVI* = 0.1)

(c) Averaged fine-scale model (*PVI* = 0.5)

(d) DL-based upscaling model (*PVI* = 0.5)

Fig. 20. Gas saturation field sample with relative error of P90.



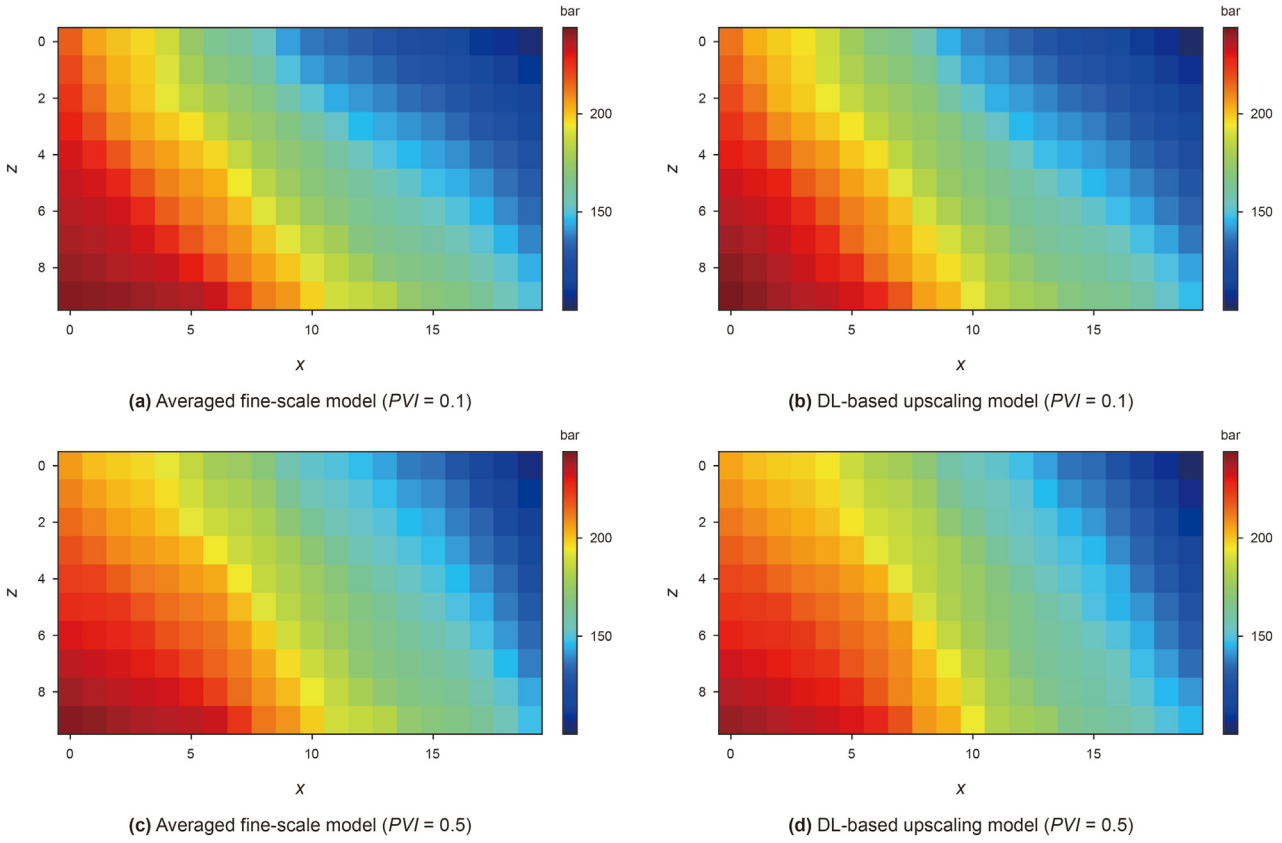
(a) Averaged fine-scale model (*PVI* = 0.1)

(b) DL-based upscaling model (*PVI* = 0.1)

(c) Averaged fine-scale model (*PVI* = 0.5)

(d) DL-based upscaling model (*PVI* = 0.5)

Fig. 21. Gas saturation field sample with relative error of P50.



(a) Averaged fine-scale model (*PVI* = 0.1)

(b) DL-based upscaling model (*PVI* = 0.1)

(c) Averaged fine-scale model (*PVI* = 0.5)

(d) DL-based upscaling model (*PVI* = 0.5)

Fig. 22. Pressure field sample with relative error of P90.

Table 2
Performance metrics for surrogate models on training and validation sets.

Surrogate model	Error metric	Training set	Validation set
*WI**	R^2	1.0000	1.0000
	MSE	0.0009	0.0008
T_x^*	R^2	0.9997	0.9996
	MSE	0.0033	0.0034
T_z^*	R^2	0.9998	0.9997
	MSE	0.0040	0.0044
P_c^*	R^2	1.0000	1.0000
	MSE	0.0001	0.0001
$K_{rw,x}^*$	R^2	0.9621	0.9613
	MSE	0.5562	0.5677
$K_{rg,x}^*$	R^2	0.9897	0.9895
	MSE	0.1595	0.1621
$K_{rw,z}^*$	R^2	0.9751	0.9748
	MSE	0.3542	0.3575
$K_{rg,z}^*$	R^2	0.9935	0.9935
	MSE	0.0923	0.0924

We further analyze the accuracy of the flow rates provided by the DL-based upscaling models. Gas and water flow rate samples with relative errors of P90, P75, P50, and P25 are presented in Figs. 25 and 26. The results indicate that the DL-based upscaling models provide reasonably accurate results, even though no significant improvement is observed for water flow rates. This suggests that most models in the test set are able to achieve reliable accuracy for both gas and water phase flow rates.

Fig. 27 presents box plots comparing the relative errors of gas saturation and pressure fields between traditional and DL-based upscaling models. The elements of the box plots follow the same conventions as previously described. The DL-based upscaling models demonstrate significantly lower relative errors compared to traditional upscaling models, with notable improvements observed in both the gas saturation and the pressure fields. In addition, The DL-based upscaling models generally exhibit narrower error ranges and lower median values, further indicating their superior performance in reducing uncertainty and enhancing accuracy.

Figs. 28 and 29 show gas saturation field samples with relative errors of P90 and P50, respectively. These samples are taken from two injection stages: *PVI* = 0.1 (before gas breakthrough) and *PVI* = 0.5 (after gas breakthrough). In the test set, 90% of the models perform better than the P90 result shown in Fig. 28, while 50% outperform those presented in Fig. 29. These results demonstrate reasonable accuracy, indicating that the majority of DL-based upscaling models can produce adequately accurate gas saturation results. The pressure field results provided by DL-based upscaling models are highly accurate, as demonstrated by the P90 relative error sample shown in Fig. 30.

7. Discussion

7.1. Speedup factors

We first compare the time costs of fine- and coarse-scale models. In this study, numerical computations (i.e., numerical simulations and upscaling computations) are performed on an AMD EPYC 7543 CPU, while DL operations are executed on an NVIDIA A100 GPU. The training time for each case was approximately 2.5 h. Due to the large number of geological models (8000

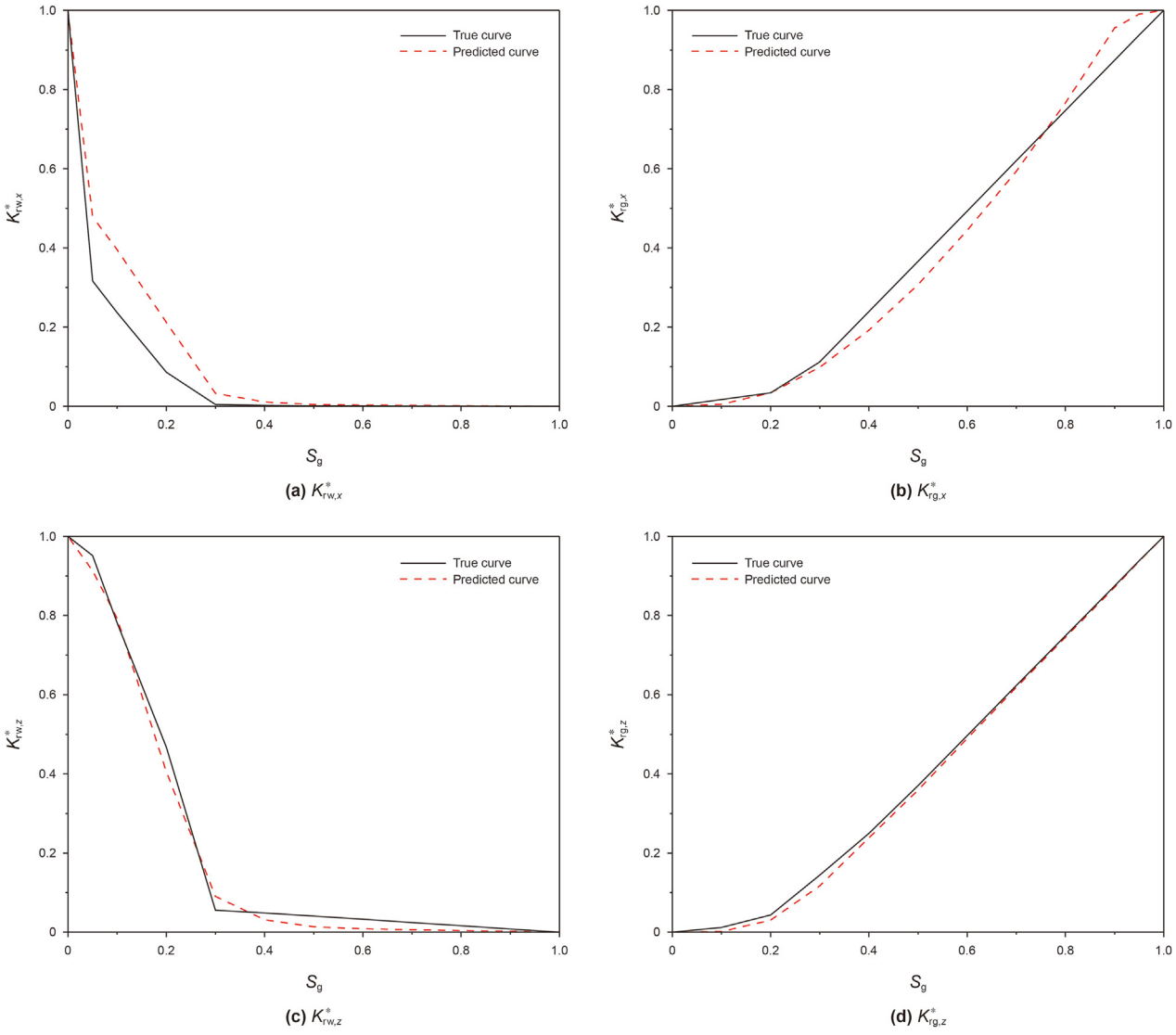


Fig. 23. Schematics showing $K^*_{rw,x}$, $K^*_{rg,x}$, $K^*_{rw,z}$, and $K^*_{rg,z}$ samples with P90 MSE error in the validation set.

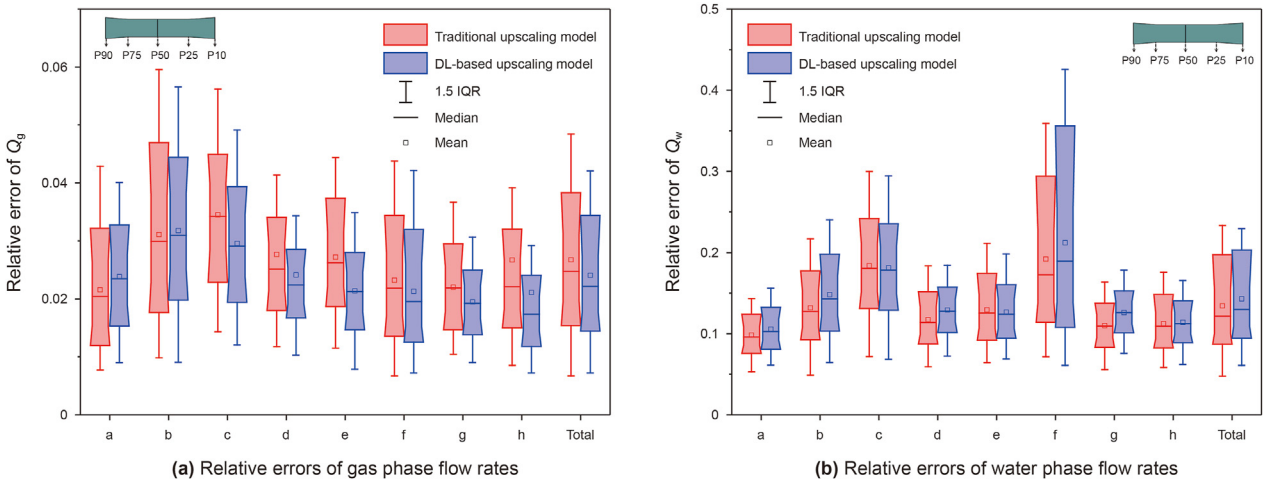


Fig. 24. Box plots showing the relative errors of gas phase and water phase flow rates at the producer for traditional and DL-based upscaling models.

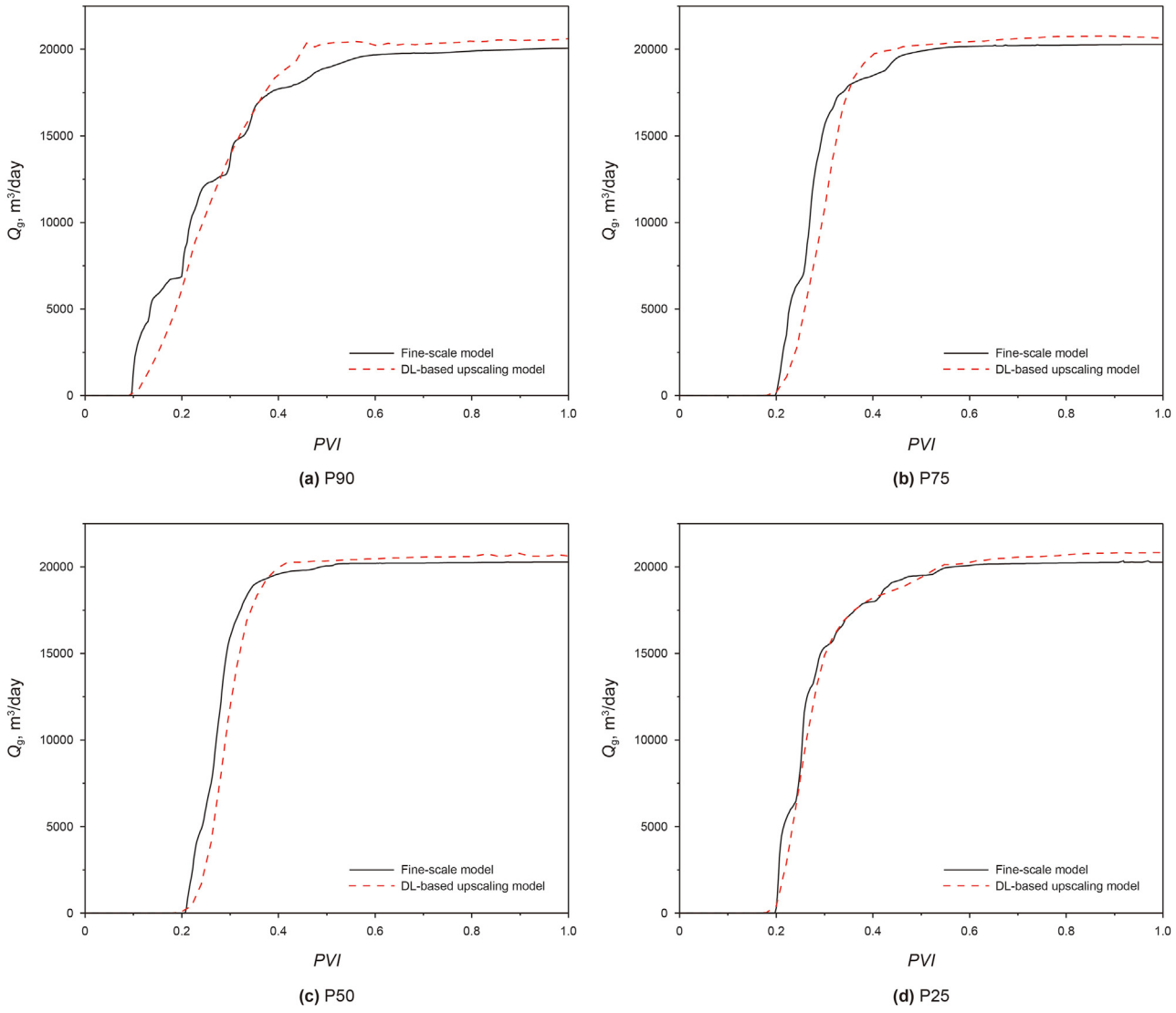


Fig. 25. Gas flow rate samples with relative errors of P90, P75, P50, and P25.

per case) and the complexities associated with simulation time (e.g., time-step cutting and Newton iteration convergence, etc.), accurately measuring the computation time for each model is challenging. Therefore, we estimate the mean computation time for a single model.

Table 3 presents the mean computation time for single fine- and coarse-scale models. The fine-scale model serves as the baseline, with a time cost of 2000 s. The traditional upscaling model includes two components: upscaling computation taking 1700 s and coarse-scale simulation taking 5 s, with a total time of 1705 s and a speedup factor of 1.17. As described in Section 2, this traditional upscaling method requires solving flow equations over the global domain, which is highly accurate but does not provide high speedup factors. Therefore, its application is often limited to scenarios that require repeated use of coarse-scale models (e.g., history matching or optimization). In contrast, our DL-based upscaling model demonstrates significant efficiency, with a time cost of only 1.5 s for predicting all the upscaling parameters, approximately 1133 times faster than traditional upscaling computation. The model achieves a total time cost of 6.5 s and an impressive speedup factor of 307.69.

7.2. Accuracy improvement

Next, we further discuss the accuracy improvement achieved by DL surrogate models. As we know, DL surrogate models typically introduce some degree of error due to their inherent nature. However, our DL-based upscaling models are able to demonstrate improved accuracy in simulation results compared to the traditional upscaling methods. This improvement occurs because the surrogate models are trained using local information, and a considerable portion of inaccurate upscaled parameters are eliminated through preprocessing methods (detailed in Section 5.2). As a result, we construct high-quality datasets that further enhance the capabilities of the surrogate models. These surrogate models correct inaccurate data through their generalization ability, leading to improved simulation results for DL-based upscaling models.

An additional finding is that the prediction performance of the relative permeability-related surrogate models declines in the non-primary flow direction, further indicating that weak flow tends to cause inaccurate upscaled data. In traditional upscaling procedures, inaccurate upscaled parameters are usually not feasible to address. Although replacement data are calculated to replace some

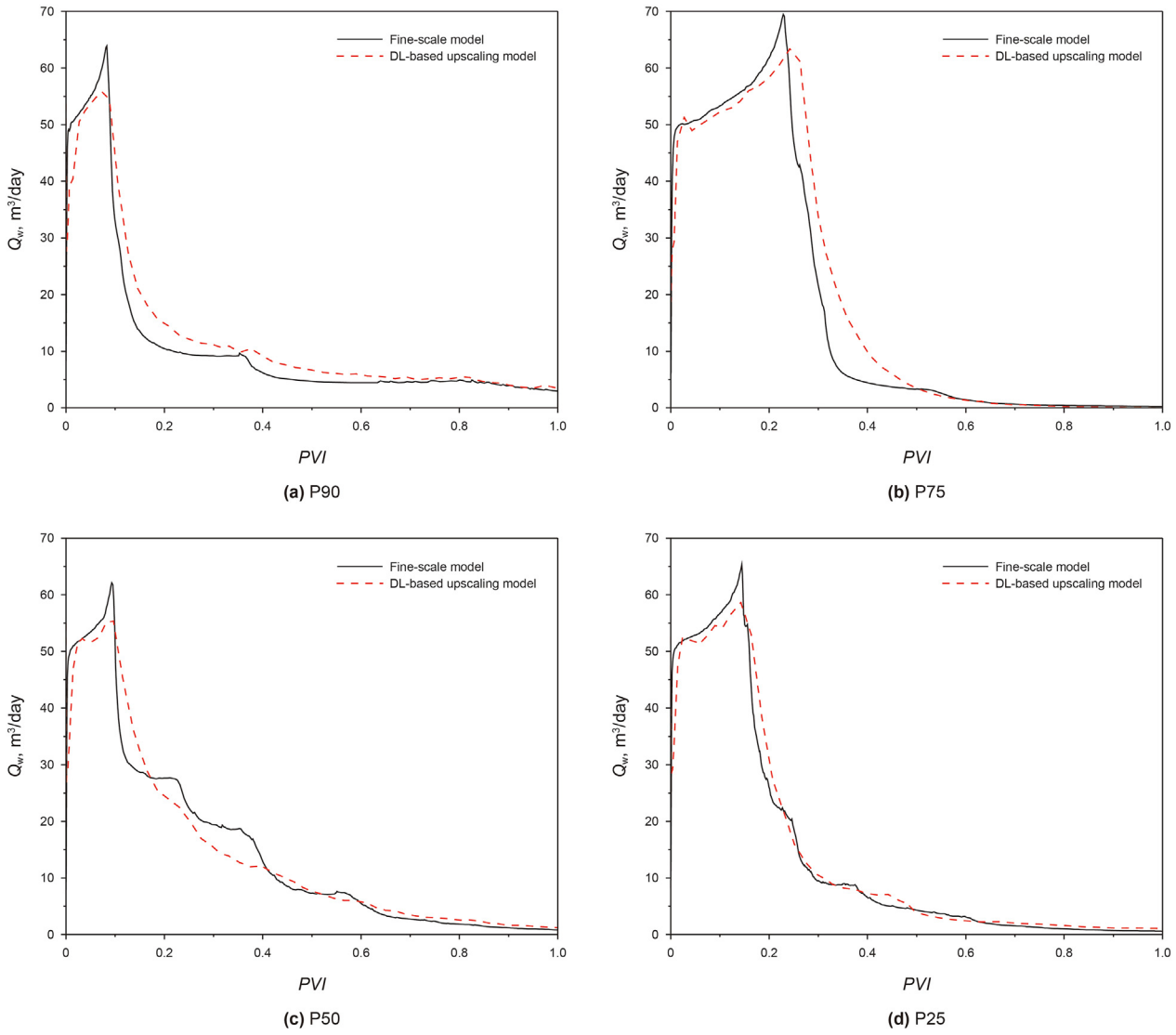


Fig. 26. Water flow rate samples with relative errors of P90, P75, P50, and P25.

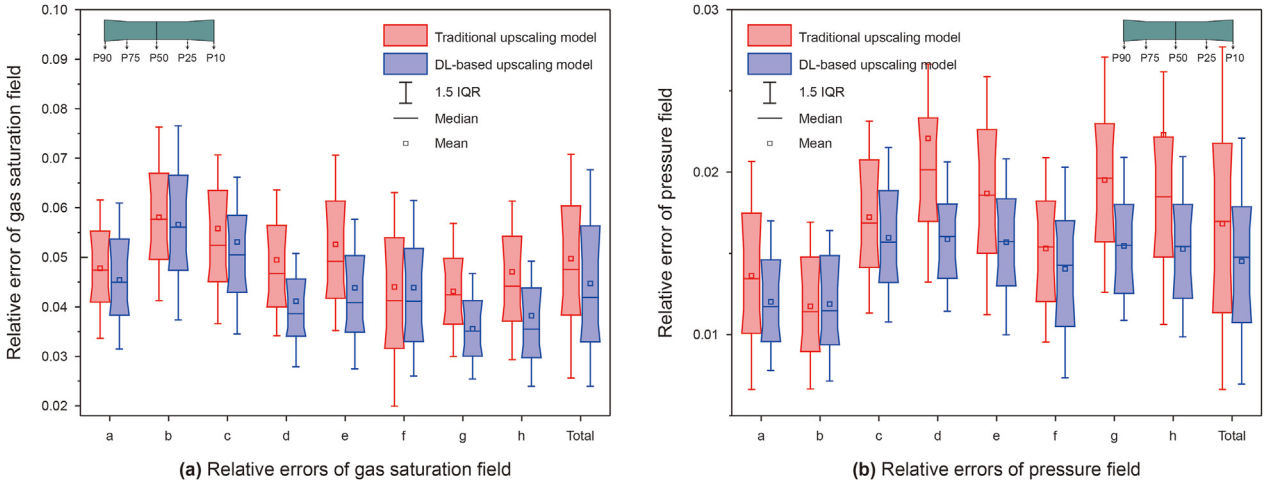
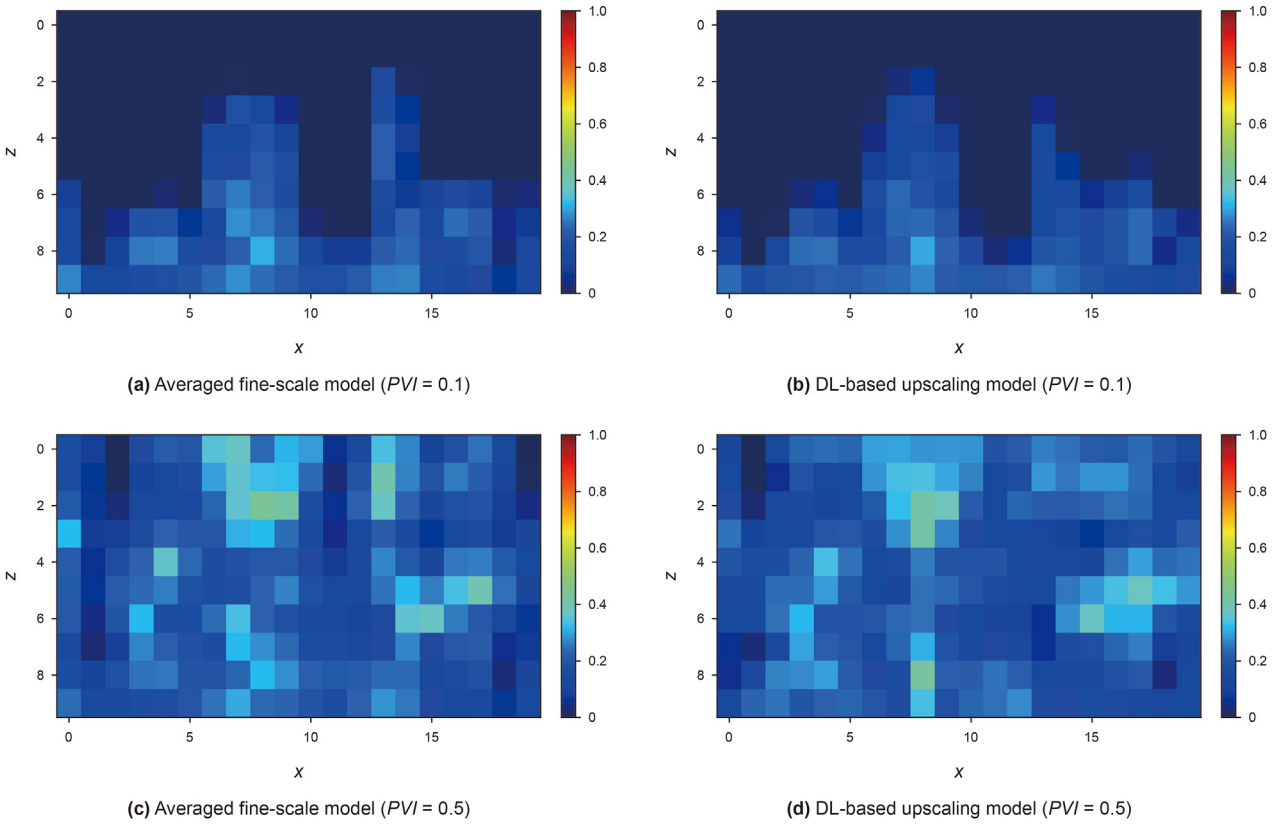


Fig. 27. Box plots showing the relative errors of gas saturation and pressure fields for traditional and DL-based upscaling models.



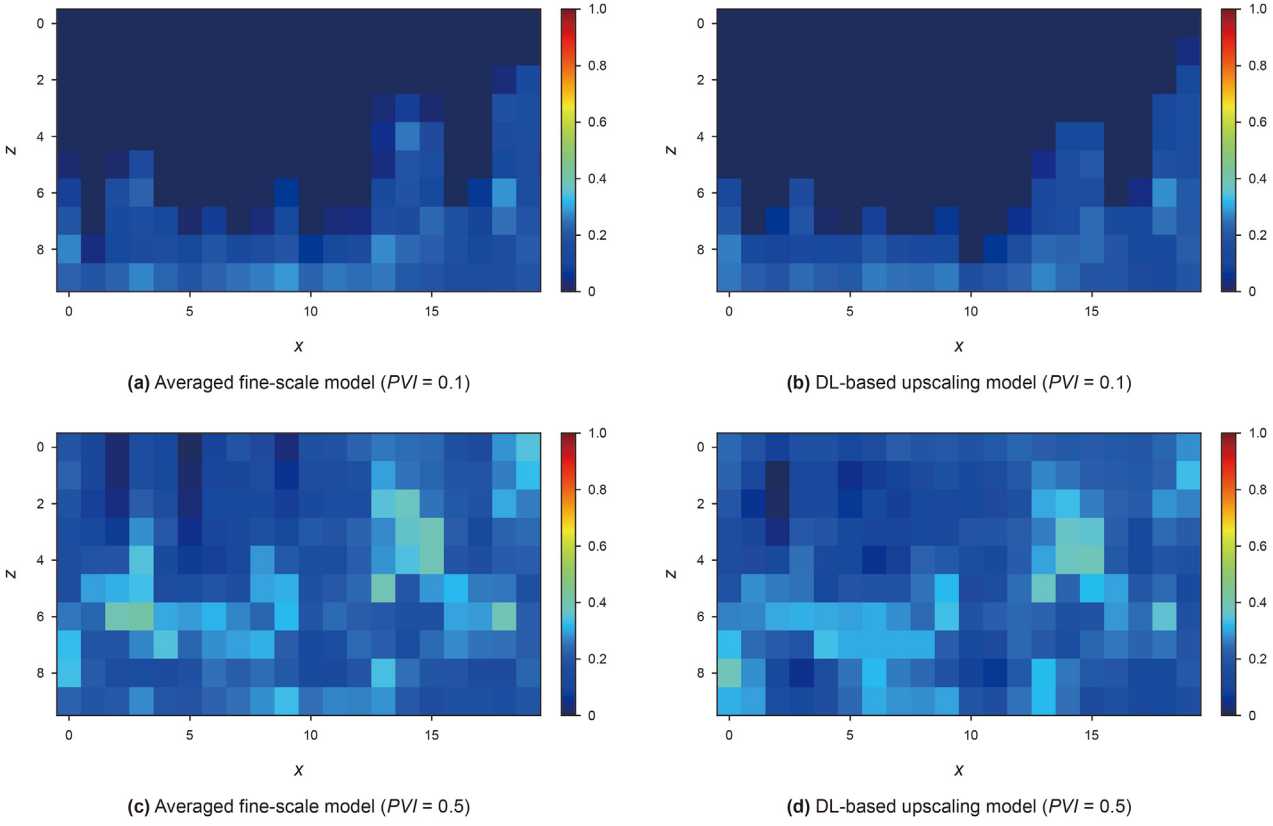
(a) Averaged fine-scale model (*PVI* = 0.1)

(b) DL-based upscaling model (*PVI* = 0.1)

(c) Averaged fine-scale model (*PVI* = 0.5)

(d) DL-based upscaling model (*PVI* = 0.5)

Fig. 28. Gas saturation field sample with relative error of P90.



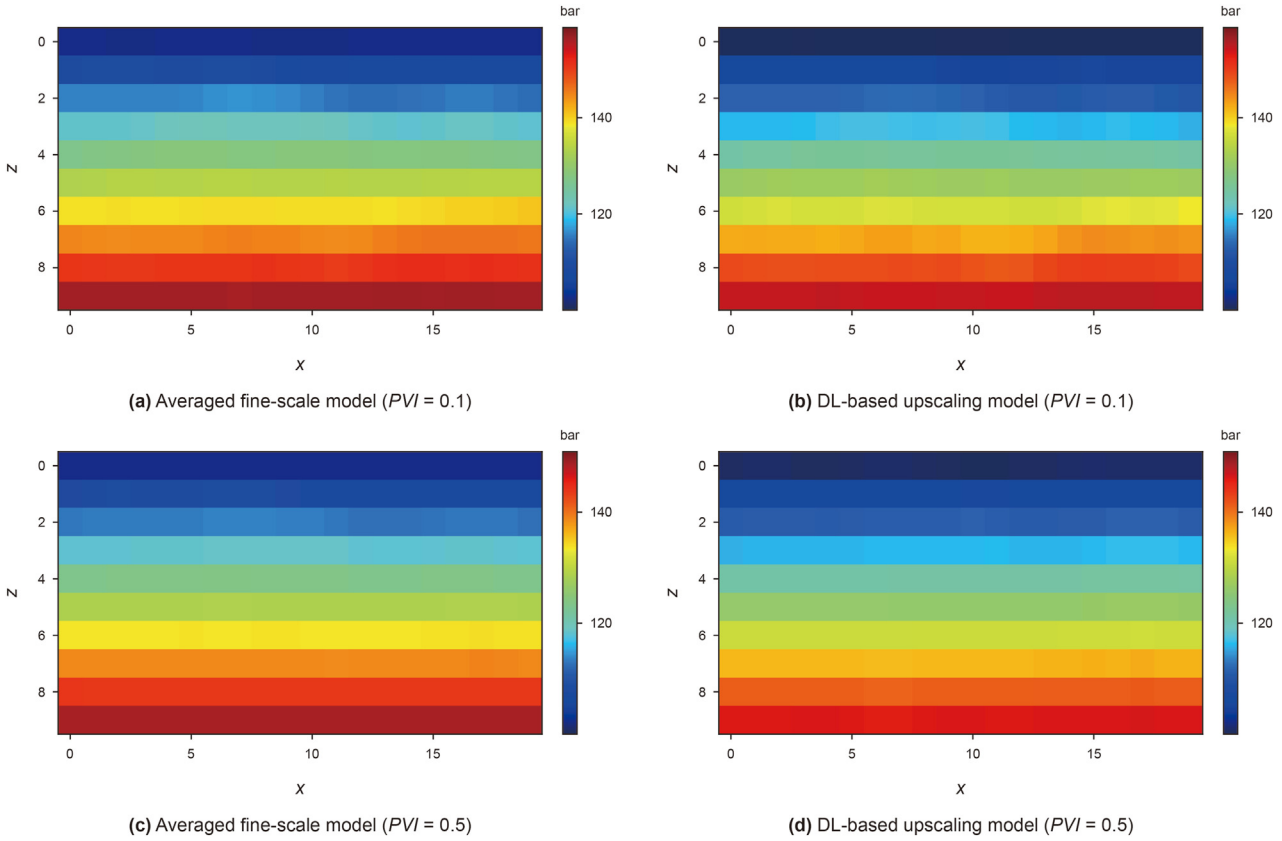
(a) Averaged fine-scale model (*PVI* = 0.1)

(b) DL-based upscaling model (*PVI* = 0.1)

(c) Averaged fine-scale model (*PVI* = 0.5)

(d) DL-based upscaling model (*PVI* = 0.5)

Fig. 29. Gas saturation field sample with relative error of P50.



(a) Averaged fine-scale model (*PVI* = 0.1)

(b) DL-based upscaling model (*PVI* = 0.1)

(c) Averaged fine-scale model (*PVI* = 0.5)

(d) DL-based upscaling model (*PVI* = 0.5)

Fig. 30. Pressure field sample with relative error of P90.

Table 3
Mean computation time for single fine- and coarse-scale models.

Model		Time cost, s	Total time cost, s	Speedup factor
Fine-scale model	Fine-scale simulation	2000	2000	–
Traditional upscaling model	Upscaling computation Coarse-scale simulation	1700 5	1705	1.17
DL-based upscaling model	DL-based upscaling Corse-scale simulation	1.5 5	6.5	307.69

inaccurate (i.e., unphysical) upscaled parameters, as described in Section 2.2, these replacement data are still inaccurate. Our DL methods offer a potential solution for improving inaccurate data generated by flow-based upscaling.

8. Conclusions and future work

In this paper, we develop a DL-based upscaling method for CO_2 injection into saline aquifers. We consider full-physics geological models that incorporate mutual solubility of CO_2 and brine, salt precipitation, relative permeability hysteresis, gravity, compressibility of rock and fluids, and capillary heterogeneity. The upscaling procedure consists of global flow-based methods for transmissibility, well index, and relative permeability, along with a steady-state method for capillary pressure. Surrogate models are constructed for these upscaled parameters by integrating CNN, Transformer encoder, and FNO architectures. These surrogate models are trained using local information, with preprocessing procedures to address inaccurate upscaled parameters generated during the flow-based upscaling processes.

Eight different types of permeability fields including 8000 models were generated. Two cases were presented using these models with flow patterns of left-to-right and bottom-to-top, respectively. In each case, 40% of the models were allocated to the training set, 10% to the validation set, and 50% to the test set. After the training process, the validation set is used to evaluate the DL predictions. The results indicate that the surrogate models can accurately predict most upscaled parameters (i.e., upscaled transmissibility, well index, capillary pressure, and relative permeabilities in primary flow direction). However, the prediction of upscaled relative permeabilities in the non-primary flow direction is less accurate due to noise data caused by weak flows. Nevertheless, the error in the non-primary flow direction has a limited impact on the final simulation results.

The trained surrogate models were then employed to perform upscaling on the test set. A key finding is that the DL-based upscaling models can improve the accuracy of simulation results compared to traditional upscaling models. This improvement

occurs because the surrogate models are trained using a high-quality training set constructed through our preprocessing procedure, making their predictions more accurate than the inaccurate upscaled parameters generated by traditional upscaling. Additionally, as expected, the DL-based upscaling achieves significant speedup factors, with upscaling accelerating approximately 1133 times compared to traditional upscaling, and the total time accelerating by 307.69 times compared to fine-scale simulation.

There are still several topics that should be considered in future research. First, the current study concentrates on the CO_2 injection stage. Future work should further incorporate post-injection processes, such as residual gas trapping and mineral trapping, into the upscaling method and DL models. Second, although two-dimensional geological models with isotropic rock properties and constant porosity are employed for computational efficiency, they do not adequately represent subsurface formations. Future work should expand both the geological models and the DL-based upscaling approach to more realistic three-dimensional scenarios. Third, to simplify the problem, this study does not incorporate a comprehensive range of parameters in the geological models, such as reservoir conditions, geological model types, rock properties, and well configurations. In future work, a more comprehensive set of such parameters should be considered to enhance the generalization ability of DL models. Finally, the application of the DL-based upscaling method to realistic cases should be pursued.

CRediT authorship contribution statement

Yan-Ji Wang: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Yan Jin:** Writing – review & editing. **Bo-Tao Lin:** Writing – review & editing. **Hui-Wen Pang:** 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.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (Nos. 52334001, U24B200085) and the Science Foundation of China University of Petroleum, Beijing (No. 2462025XKBH012). Yan-Ji Wang is grateful for Prof. Louis Durlofsky's guidance during his time as a visiting student researcher in the Department of Energy Science and Engineering at Stanford University. We would also like to thank Stanford University for providing the computational resources, and Dr. Su Jiang for discussion about deep learning methods.

References

Behzadi, H., Alvarado, V., 2012. Upscaling of upward CO_2 migration in 2D system. Adv. Water Resour. 46, 46–54. https://doi.org/10.1016/j.advwatres.2012.06.003.

Bohne, R., 2018. Machine-learning Algorithms for the Computation of Upscaled Permeabilities. MS Thesis. NTNU.

Bouquet, S., Bruel, D., de Fouquet, C., 2013. Influence of heterogeneities and upscaling on CO_2 storage prediction at large scale in deep saline aquifer. Energy Proc. 37, 4445–4456. https://doi.org/10.1016/j.egypro.2013.06.349.

Cameron, D.A., Durlofsky, L.J., 2012. Optimization of well placement, CO_2 injection rates, and brine cycling for geological carbon sequestration. Int. J. Greenh. Gas Control 10, 100–112. https://doi.org/10.1016/j.ijggc.2012.06.003.

Chen, Y., 2005. Upscaling and Subgrid Modeling of Flow and Transport in Heterogeneous Reservoirs. Ph.D Dissertation. Stanford University.

Chen, Y., Durlofsky, L.J., Gerritsen, M., Wen, X.H., 2003. A coupled local–global upscaling approach for simulating flow in highly heterogeneous formations. Adv. Water Resour. 26 (10), 1041–1060. https://doi.org/10.1016/S0309-1708(03)00101-5.

Cheng, K.B., Rabinovich, A., 2020. Upscaling capillary pressure curves for numerical modeling of gravity-capillary driven flow. Adv. Water Resour. 142, 103639. https://doi.org/10.1016/j.advwatres.2020.103639.

Deutsch, C.V., Journel, A.G., 1992. Geostatistical Software Library and User's Guide. N. Y. 119 (147).

Durlofsky, L.J., 2005. Upscaling and gridding of fine scale geological models for flow simulation. In: 8th International Forum on Reservoir Simulation.

Efendiev, Y.R., Durlofsky, L.J., 2004. Accurate subgrid models for two-phase flow in heterogeneous reservoirs. SPE J. 9 (2), 219–226. https://doi.org/10.2118/88363-PA.

Han, Y., Hamon, F.P., Jiang, S., Durlofsky, L.J., 2024. Surrogate model for geological CO_2 storage and its use in hierarchical MCMC history matching. Adv. Water Resour. 187, 104678. https://doi.org/10.1016/j.advwatres.2024.104678.

Ide, S.T., Jessen, K., Orr Jr, F.M., 2007. Storage of CO_2 in saline aquifers: effects of gravity, viscous, and capillary forces on amount and timing of trapping. Int. J. Greenh. Gas Control 1 (4), 481–491. https://doi.org/10.1016/S1750-5836(07)00091-6.

Itthisawatpan, K., 2013. Upscaling of Two-phase Flow with Capillary Pressure Heterogeneity Effects. MS Thesis. Stanford University.

Jo, H., Pan, W., Santos, J.E., Jung, H., Pyrcz, M.J., 2021. Machine learning assisted history matching for a deepwater lobe system. J. Petrol. Sci. Eng. 207, 109086. https://doi.org/10.1016/j.petrol.2021.109086.

Juanes, R., MacMinn, C., 2008. Upscaling of capillary trapping under gravity override: application to CO_2 sequestration in aquifers. In: SPE Improved Oil Recovery Conference. https://doi.org/10.2118/113496-MS.

Kingma, D.P., 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kou, Z., Wang, H., Alvarado, V., McLaughlin, J.F., Quillinan, S.A., 2022. Method for upscaling of CO_2 migration in 3D heterogeneous geological models. J. Hydrol. 613, 128361. https://doi.org/10.1016/j.jhydrol.2022.128361.

Kwon, S., Park, G., Jang, Y., Cho, J., Chu, M.G., Min, B., 2021. Determination of oil well placement using convolutional neural network coupled with robust optimization under geological uncertainty. J. Petrol. Sci. Eng. 201, 108118. https://doi.org/10.1016/j.petrol.2020.108118.

Leverett, M., 1941. Capillary behavior in porous solids. Transactions of the AIME 142 (1), 152–169. https://doi.org/10.2118/941152-G.

Li, H., Durlofsky, L.J., 2016. Upscaling for compositional reservoir simulation. SPE J. 21 (3), 873–887. https://doi.org/10.2118/173212-PA.

Li, H., Misra, S., 2021. Reinforcement learning based automated history matching for improved hydrocarbon production forecast. Appl. Energy 284, 116311. https://doi.org/10.1016/j.apenergy.2020.116311.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., Anandkumar, A., 2020. Fourier neural operator for parametric partial differential equations. arXiv preprint arXiv:2010.08895.

Li, Z., Huang, D.Z., Liu, B., Anandkumar, A., 2023. Fourier neural operator with learned deformations for pdes on general geometries. J. Mach. Learn. Res. 24 (338), 1–26.

Liu, F.T., Ting, K.M., Zhou, Z.H., 2008. Isolation forest. In: 8th IEEE International Conference on Data Mining. https://doi.org/10.1109/ICDM.2008.17.

Lohne, A., Virnovsky, G., 2006. Three-phase upscaling in capillary and viscous limit. In: SPE Improved Oil Recovery Conference. https://doi.org/10.2118/99567-MS.

Madasu, S., Siddiqui, S., Rangarajan, K.P., 2019. Compressing time-dependent reservoir simulations using graph-convolutional neural network G-CNN. In: Abu Dhabi International Petroleum Exhibition and Conference. https://doi.org/10.2118/197444-MS.

Mouche, E., Hayek, M., Mügler, C., 2010. Upscaling of CO_2 vertical migration through a periodic layered porous medium: the capillary-free and capillary-dominant cases. Adv. Water Resour. 33 (9), 1164–1175. https://doi.org/10.1016/j.advwatres.2010.07.005.

Rabinovich, A., Itthisawatpan, K., Durlofsky, L.J., 2015. Upscaling of CO_2 injection into brine with capillary heterogeneity effects. J. Petrol. Sci. Eng. 134, 60–75. https://doi.org/10.1016/j.petrol.2015.07.021.

Rabinovich, A., Li, B., Durlofsky, L.J., 2016. Analytical approximations for effective relative permeability in the capillary limit. Water Resour. Res. 52 (10), 7645–7667. https://doi.org/10.1002/2016WR019234.

Remy, N., Boucher, A., Wu, J., Alexandre, B., 2009. Applied Geostatistics with SGeMS: A User's Guide. Cambridge University Press.

Renard, P., De Marsily, G., 1997. Calculating equivalent permeability: a review. Adv. Water Resour. 20 (5–6), 253–278. https://doi.org/10.1016/S0309-1708(96)00050-4.

Romeu, R., Noetinger, B., 1995. Calculation of internodal transmissivities in finite difference models of flow in heterogeneous porous media. Water Resour. Res. 31 (4), 943–959. https://doi.org/10.1029/94WR02422.

Saadatpoor, E., Bryant, S.L., Sepehrnoori, K., 2011. Effect of upscaling heterogeneous domain on CO_2 trapping mechanisms. Energy Proc. 4, 5066–5073. https://doi.org/10.1016/j.egypro.2011.02.480.

Selvadurai, P., Selvadurai, A., 2014. On the effective permeability of a heterogeneous porous medium: the role of the geometric mean. Philos. Mag. 94 (20), 2318–2338. https://doi.org/10.1080/14786435.2014.913111.

Spycher, N., Pruess, K., 2005. CO_2-H_2O mixtures in the geological sequestration of CO_2. II. Partitioning in chloride brines at 12–100 °C and up to 600 bar. Geochem. Cosmochim. Acta 69 (13), 3309–3320. https://doi.org/10.1016/j.gca.2005.01.015.

Spycher, N., Pruess, K., Ennis-King, J., 2003. CO_2-H_2O mixtures in the geological

sequestration of CO_2. I. Assessment and calculation of mutual solubilities from 12 to 100 °C and up to 600 bar. Geochem. Cosmochim. Acta 67 (16), 3015–3031. https://doi.org/10.1016/S0016-7037(03)00273-4.

Srinivasan, S., O'Malley, D., Mudunuru, M.K., Sweeney, M.R., Hyman, J.D., Karra, S., Frash, L., Carey, J.W., Gross, M.R., Guthrie, G.D., 2021. A machine learning framework for rapid forecasting and history matching in unconventional reservoirs. Sci. Rep. 11 (1), 21730. https://doi.org/10.1038/s41598-021-01023-w.

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. Math. Geol. 34 (1), 1–21. https://doi.org/10.1023/A:1014009426274.

Strebelle, S.B., 2000. Sequential Simulation Drawing Structures from Training Images. Ph.D Dissertation. Stanford University.

Tang, H., Durlofsky, L.J., 2022. Use of low-fidelity models with machine-learning error correction for well placement optimization. Comput. Geosci. 26 (5), 1189–1206. https://doi.org/10.1007/s10596-022-10153-7.

Tang, L., Li, J., Lu, W., Lian, P., Wang, H., Jiang, H., Wang, F., Jia, H., 2021. Well control optimization of waterflooding oilfield based on deep neural network. Geofluids 2021 (1), 8873782. https://doi.org/10.1155/2021/8873782.

Wang, N., Chang, H., Kong, X.-Z., Zhang, D., 2023a. Deep learning based closed-loop well control optimization of geothermal reservoir with uncertain permeability. Renew. Energy 211, 379–394. https://doi.org/10.1016/j.renene.2023.04.088.

Wang, N., Liao, Q., Chang, H., Zhang, D., 2023b. Deep-learning-based upscaling method for geologic models via theory-guided convolutional neural network. Comput. Geosci. 27 (6), 913–938. https://doi.org/10.1007/s10596-023-10233-2.

Wang, Y., Li, H., Xu, J., Liu, S., Wang, X., 2022. Machine learning assisted relative permeability upscaling for uncertainty quantification. Energy 245, 123284. https://doi.org/10.1016/j.energy.2022.123284.

Wang, Y., Li, H., Xu, J., Liu, S., Tan, Q., Wang, X., 2023. Machine learning assisted two-phase upscaling for large-scale oil-water system. Appl. Energy 337, 120854. https://doi.org/10.1016/j.apenergy.2023.120854.

Wen, G., Hay, C., Benson, S.M., 2021. CCSNet: a deep learning modeling suite for CO_2 storage. Adv. Water Resour. 155, 104009. https://doi.org/10.1016/j.advwatres.2021.104009.

Wen, G., Li, Z., Azizzadenesheli, K., Anandkumar, A., Benson, S.M., 2022. U-FNO—an enhanced Fourier neural operator-based deep-learning model for multiphase flow. Adv. Water Resour. 163, 104180. https://doi.org/10.1016/j.advwatres.2022.104180.

Wen, G., Li, Z., Long, Q., Azizzadenesheli, K., Anandkumar, A., Benson, S.M., 2023. Real-time high-resolution CO_2 geological storage prediction using nested Fourier neural operators. Energy Environ. Sci. 16 (4), 1732–1741. https://doi.org/10.1039/D2EE04204E.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., 2020. Transformers: state-of-the-art natural language processing. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. https://doi.org/10.18653/v1/2020.emnlp-demos.6.

Zhang, X., Ma, F., Yin, S., Wallace, C.D., Soltanian, M.R., Dai, Z., Ritzi, R.W., Ma, Z., Zhan, C., Lü, X., 2021. Application of upscaling methods for fluid flow and mass transport in multi-scale heterogeneous media: a critical review. Appl. Energy 303, 117603. https://doi.org/10.1016/j.apenergy.2021.117603.

Zou, A., Durlofsky, L.J., 2023. Integrated framework for constrained optimization of horizontal/deviated well placement and control for geological CO_2 storage. SPE J. 28 (5), 2462–2481. https://doi.org/10.2118/212228-PA.