

Original Paper

Deep learning-based upscaling for reservoir models on corner-point grids via finite-volume physics-informed Fourier neural operator

Jing-Qi Lin^{a,b}, Xia Yan^{a,b,*}, Kai Zhang^{a,c}, Qi Zhang^d, Zhao Zhang^e, Li-Ming Zhang^{a,b}, Pi-Yang Liu^c

^a School of Petroleum Engineering, China University of Petroleum (East China), Qingdao, 266580, Shandong, China

^b National Key Laboratory of Deep Oil and Gas, China University of Petroleum (East China), Qingdao, 266580, Shandong, China

^c Civil Engineering School, Qingdao University of Technology, Qingdao, 266520, Shandong, China

^d Department of Civil and Environmental Engineering, University of Macau, Taipa, 999078, Macao SAR, China

^e Research Centre for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, 266237, Shandong, China

ARTICLE INFO

Article history:

Received 21 October 2025

Received in revised form

21 January 2026

Accepted 21 April 2026

Available online 27 April 2026

Edited by Meng-Jiao Zhou

Keywords:

Physics-informed

Fourier neural operator

Parametric learning

Corner-point grid model

Flow-based upscaling

ABSTRACT

This study proposes a neural network (NN)-assisted, physics-informed framework for global flow-based upscaling to accelerate reservoir simulation. While traditional global upscaling methods can accurately capture fine-scale flow characteristics, their high computational cost hinders practical application in ensemble-based uncertainty quantification and optimization workflows. Existing NN-assisted upscaling approaches predominantly rely on numerical simulation data, which remains computationally expensive and often lacks physical interpretability. Furthermore, no current method enables physics-informed training for corner-point grid (CPG) models without simulation data, limiting the application of deep learning in practical geological model upscaling. To address these limitations, we develop a finite-volume physics-informed Fourier neural operator (FV-PIFNO) as a pressure-solution surrogate model for CPG models. This physics-driven approach eliminates the need for simulation data and inherently enforces inter-grid flux continuity during inference. A flow-based numerical post-processing procedure is designed to compute coarse-grid transmissibility and well index that are strictly equivalent to fine-scale fluxes, naturally extending to non-orthogonal CPG systems. Comparisons are designed between different mapping and driving methods. Validation on synthetic models demonstrates that the fully physics-constrained approach achieves high accuracy in both single-phase and two-phase flow simulations, along with strong generalization to heterogeneous scenarios. Application to a standard SAIGUP model confirms that the framework achieves accuracy comparable to numerical upscaling while significantly improving computational efficiency, demonstrating its potential for practical, high-fidelity reservoir upscaling.

© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Groundwater contamination (Neukum and Azzam, 2009; Banaei et al., 2021), hydrocarbon reservoir development (Yan et al., 2019; Wu, 2016), geological storage of carbon dioxide (Zhang and Agarwal, 2013; Kumar et al., 2005), and underground hydrogen storage (Zhao et al., 2024; Feldmann et al., 2016) require dynamic analysis, scheme design, and decision-making that

depend on numerical simulation for prediction and evaluation. Geologists and engineers usually employ fine grids to model complex subsurface structures (Houlding, 1994). Although these high-resolution models can provide detailed descriptions of geological features, directly applying them to fluid-flow simulations imposes substantial computational costs. Therefore, model upscaling is frequently required (Farmer, 2002). The key physical properties of fine-grid models are reasonably homogenized to a

* Corresponding author.

E-mail addresses: jsyanxia1989@163.com, 20190025@upc.edu.cn (X. Yan).

Peer review under the responsibility of China University of Petroleum (Beijing).

coarser scale, thereby greatly reducing computational time while maintaining comparable simulation accuracy. Permeability, as one of the critical non-additive parameters characterizing subsurface fluid-flow capacity, strongly influences the approximation for the original fine model (Wu et al., 2002). An ideal upscaled model should generate reasonable coarse-scale parameters that the simulation results remain consistent with the fine-scale accurate solutions (Qi and Hesketh, 2005; Krishnamurthy et al., 2024). Consequently, developing upscaling methods that combine accuracy, efficiency, and applicability has become a popular and challenging direction.

Over the past decades, various upscaling methods have been proposed, which can be categorized into analytical approaches and numerical-solution-based approaches (Qi and Hesketh, 2005). Among them, flow-based methods that rely on numerical solutions are considered to have strong applicability, including purely local, extended local, global, and quasi-global types (Durlafsky, 2005). Flow-based methods typically impose boundary conditions locally or globally, solve the flow equations, and compute coarse-grid properties that are equivalent in fluid flux, thereby reproducing fine-scale characteristics as closely as possible. The results of upscaling are usually non-unique, depending on the chosen scheme, and are related to grid geometry and the imposed boundary conditions (Dagan et al., 2013). In particular, global flow-based methods impose global boundary conditions over the entire domain (White and Horne, 1987; Holden and Nielsen, 2000) and are regarded capable of providing accurate upscaling results under given well and boundary conditions (Zhang et al., 2008; Alpak, 2021). Moreover, global methods are well suited for models with complex geological structures that are modeled using corner-point grids (CPGs) (Islam et al., 2021). However, these methods usually require solving fine-grid models, which is computationally expensive. This limits their application in practical downstream scenarios, especially in simulation and optimization workflows that involve multiple numerical models. In robust optimization workflows that account for geological uncertainty (Pan et al., 2015; Siraj et al., 2016), it is often necessary to simulate ensembles of models with different permeability fields to construct parameter spaces for uncertainty assessment. In closed-loop reservoir management (CLRM) (Wang et al., 2009; Jansen et al., 2009), permeability fields are frequently updated based on collected real-time data. Some multiscale methods involve the use of multiple numerical models with different resolutions, including sequential use of models to accelerate simulations (Kostakis et al., 2020), or employing results from different simulations for multi-fidelity neural network (NN) training (Jiang and Durlafsky, 2023). Ensemble-based upscaling has also been proposed (Chen and Durlafsky, 2008), in which upscaling is performed at the ensemble level to control and correct upscaling errors. In more complex application scenarios that consider multiple factors and require rolling updates (Salehian et al., 2021; Chen et al., 2009), the number of models requiring upscaling grows in proportion to the product of multiple factors. Therefore, incorporating global methods into practical workflows results in significant computational costs, which diminishes their practical value.

In recent years, artificial intelligence technologies, with deep learning (LeCun et al., 2015) as a prominent representative, have developed rapidly. A growing number of studies have begun to exploit the strong nonlinear learning capacity of NNs to assist in model upscaling. For example, Sayyafzadeh and Guérillot (2022) treated the upscaling problem as an image downsampling task with NN assistance. Pal et al. (2023) attempted to enable NNs to learn upscaling mechanisms from extensive geological practice. (Ma et al., 2024; Kodama et al., 2024) generated large datasets of “original properties–upscaled results” through simulations for NN

training, thereby transforming the upscaling into a rapid forward inference. Ramos et al. (2023) employed optimization algorithms to refine the end-to-end surrogate NN mapping from properties to results. Santos et al. (2022) trained an NN on numerical upscaling results and then develops an adjustment procedure to enhance the inference accuracy of the NN. However, these “black-box” models contain several issues: (1) NN training requires large amounts of labeled data, the acquisition of which is computationally expensive. (2) Trained NNs usually upscale only to a fixed resolution, which limits their broader applicability. (3) The workflows lack physical constraints, so the NN outputs may be physically inconsistent.

Actually, NNs can be employed to approximate the most computationally expensive pressure-solution step in the upscaling process. A series of studies have researched numerical simulation surrogate models (NSSMs) (Tang et al., 2020; Jin et al., 2020; Lin et al., 2024). In addition, a class of deep learning methods named as neural operators has demonstrated superior performance in the parametric learning of partial differential equations (PDEs) (Lu et al., 2022). Representative methods include the deep operator network (Lu et al., 2021) and the Fourier neural operator (FNO) (Li et al., 2020). Several variants of neural operators have been proposed (Wen et al., 2022; Tariq et al., 2025), demonstrating strong capability in learning subsurface flow problems. In a recent study (Wang et al., 2025), the FNO was applied to learn the direct mapping from fine-grid parameters to upscaled results, and it exhibited strong inference capability. However, most existing studies focus on improving the accuracy of NSSMs, and research on assisting upscaling remains scarce. Moreover, as far as we know, no study has employed neural operators to construct NSSMs for the purpose of assisting model upscaling.

Concurrently, physics-informed neural networks (PINNs) (Raissi et al., 2019; Cai et al., 2021) have provided profound insights into improving the learning of physical problems by NNs. Physics-constrained approaches have also been applied to the training of NSSMs (Donnelly et al., 2024; Wang et al., 2020), including the integration with the neural operator (Li et al., 2024; Wang et al., 2021). The incorporation of physical knowledge improves inference accuracy comparing with purely data-driven approaches (Yang et al., 2024; Koric and Abueidda, 2023). Furthermore, several studies have proposed physics-constrained loss functions based on discretized PDEs (Zhang et al., 2023; Yan et al., 2024) to address the convergence difficulties of automatic differentiation (AD)-based losses in heterogeneous problems (Lin et al., 2025). Such discretized losses have also been employed in the training of NSSMs (Wang et al., 2022). Moreover, recent studies have combined discretized physical constraints with neural operator methods to further enhance model performance (Rezaei et al., 2025; Eshaghi et al., 2025; Xu et al., 2024; Harandi et al., 2025). To date, most deep learning-assisted upscaling methods still lack physical constraints (Wang et al., 2023) utilized physics-constrained NNs to construct surrogate models of the local pressure solving procedure, thereby accelerating the local flow-based upscaling. However, there is still a lack of research on physics-driven NSSMs for CPG models, as well as a lack of NN-assisted flow-based upscaling methods applicable to CPGs.

In this study, we design a physics-driven NN-assisted upscaling method considering permeability heterogeneity to achieve efficient flow-based upscaling of CPG reservoir models. Several NNs based on the FNO is trained to learn the fine-scale pressure-solving surrogate model, thereby replacing the computationally expensive numerical simulation procedure. The FNO can effectively capture universal low-frequency features and possesses cross-resolution inferential capability (Wang et al., 2020). During model training, discretized PDEs based on the finite volume method (FVM)

(Eymard et al., 2000) are used to construct a physics-constrained loss function. This method, referred to as the finite-volume physics-informed Fourier neural operator (FV-PIFNO) (Yan et al., 2025), defines physical consistency through control volumes and fluid fluxes, in contrast to computing derivatives at spatial sampling points via automatic differentiation. This discretization reduces the convergence difficulties caused by high-order derivatives and suppresses the amplification of high-frequency errors. Furthermore, the finite-volume formulation naturally guarantees strict local mass conservation at the discrete level for each control volume and each time step (Eymard et al., 2000), which is a fundamental requirement for strongly nonlinear multiphase flow in reservoir simulation (Chen et al., 2006). It also exhibits superior numerical robustness when applied to corner-point grids, whose geometries are highly distorted while their topological structures remain regular, thereby enabling stable treatment of complex engineering features such as fault displacements, layer pinch-outs, and well-grid coupling (Darlowsky, 2005). In contrast, weak-form finite-element discretizations do not inherently satisfy local conservation (White et al., 2021) and, when applied to corner-point grids, are prone to degraded isoparametric mappings, reduced quadrature accuracy, and ill-conditioned system matrices (Prabhune and Suresh, 2024). These issues render numerical stability highly sensitive to grid quality and significantly increase implementation complexity. The integration of FNO with FVM in this study is therefore specifically intended to inherit these advantages. FVM strictly relies on neighbor connectivity to calculate fluxes, and although CPGs exhibit complex geometries in the physical domain, they retain structured characteristics in the computational (logical) domain; FNO is inherently well-suited for such structured tensor data, capable of preserving grid topological information along with spatial global and local properties. The trained model is incorporated into the upscaling workflow, followed by explicit procedure for flux-equivalence calculation. This NN-assisted workflow can rapidly output upscaled transmissibility and well index at arbitrary coarse resolutions for CPG models. Essentially, this method embeds physical consistency into the entire chain of NN-assisted upscaling. The discretized constraint enforces flux conservation across different grids and enhances the NN inference consistency with physical laws. The coarse-scale reconstruction of transmissibility and well index strictly follows the flux equivalence and is inherently robust to the non-orthogonal connections of CPGs. To the best of our knowledge, no prior study has employed NNs to establish NSSMs for CPG reservoir models, nor to perform NN-assisted flow-based upscaling of these models. These challenges include the high cost of fine-scale simulations, grid non-orthogonality, and the low generalization capability of models in unseen domains.

The key contributions of this study are summarized as follows: (1) Data-free physics-driven workflow: We introduce a label-free training mechanism based on discretized FVM residuals, which ensures physical consistency and seamlessly couples pressure prediction with the explicit derivation of coarse-scale parameters. (2) Corner-point grid adaptability: By leveraging the logical topology of non-orthogonal CPGs, the method successfully implements physics-informed constraints on complex geometries, bridging the gap between neural operators and reservoir model grids. (3) Arbitrary coarse-scale upscaling: The model learns the fine-scale pressure mapping, enabling the explicit, post-processing derivation of upscaled parameters at arbitrary coarse resolutions without retraining.

In the case studies, we design comparisons among three representative NN-assisted upscaling pathways: fully end-to-end, data-driven NSSM-assist, and physics-driven NSSM-assist. The results indicate that the proposed physics-driven method achieves

the best upscaling accuracy and exhibits superior extrapolation capability to heterogeneous-distribution scenarios. We also conduct tests on the SAIGUP model (Howell et al., 2008) with more complex geological structures, where the new method achieves upscaling accuracy comparable to simulation methods while providing substantially improved efficiency.

The remaining content of this paper is arranged as follows. The Methodology section introduces the flow governing equations, the upscaling theory, and the neural NN-assisted upscaling methods. The Results and Discussion sections present comparisons of different case studies. The Conclusion section summarizes the main findings of this study and outlines future directions.

2. Methodology

2.1. Flow-based upscaling of reservoir models

The primary objective of flow-based upscaling of reservoir models is to approximate the effects of fine-scale heterogeneity on fluid flow behavior as closely as possible in subsequent calculations of the coarse-scale model (Darlowsky, 2005). The flow problem is typically formulated as a single-phase steady-state flow based on Darcy's law, and the governing equation is

$$\mathbf{v} = -\frac{\mathbf{K}}{\mu} \nabla p \quad (1)$$

$$\nabla \cdot \mathbf{v} + q = 0 \quad (2)$$

$$p = p_{bc} \text{ on Dirichlet boundary } \Gamma_D \quad (3)$$

$$\mathbf{v} \cdot \mathbf{n} = v_{bc} \text{ on Neumann boundary } \Gamma_N \quad (4)$$

Here, $\mathbf{v}$ is the velocity; $\mathbf{K}$ is the permeability tensor; p is the pressure; q is the source/sink term; p_{bc} is the pressure on the Dirichlet boundary; $\mathbf{n}$ is the unit normal vector; and v_{bc} is the velocity on the Neumann boundary.

Flow-based upscaling of permeability typically seeks an effective permeability tensor $\mathbf{K}^*$ to reproduce the total fine-scale flux across each homogenized coarse-scale domain, satisfying the following equation:

$$\int_{\Omega} \mathbf{K} \nabla p d\vec{x} = \mathbf{K}^* \int_{\Omega} \nabla p d\vec{x} \quad (5)$$

where Ω is any homogenized coarse-scale domain and $d\vec{x}$ is the differential volume element. The upscaled permeability tensor $\mathbf{K}^*$ is generally not uniquely determined, as it depends on the flux through Ω , which in turn is determined by the different boundary conditions (Dagan et al., 2013).

2.2. Numerical simulation-assisted upscaling

Using observed reservoir boundary conditions and specified well conditions to guide upscaling is commonly termed well-drive upscaling (Zhang et al., 2008). Global transmissibility upscaling can better capture continuous non-local geologic heterogeneities, including those modeled on corner point grids (Alpak, 2021). A clearer understanding of the boundary conditions in the actual simulation allows for the calculation of more accurate equivalent parameters (Darlowsky, 2005). In this study, we focus on transmissibility upscaling for the purpose of accelerating reservoir simulation.

At the fine scale, the model boundaries are set with no-flux Neumann boundaries, and the locations and conditions for the

injection and production wells are also defined based on actual scenarios (high pressure is assigned to injection wells and low pressure to production wells). By substituting Eq. (1) into Eq. (2) and discretizing the result using the FVM, the following discrete equation is deduced

$$\sum_{j \in G_i} \left(\frac{1}{\mu} T_{ij} (p_j - p_i) \right) + q V_i = 0 \quad (6)$$

where i and j are the indices of the fine grids; μ is the fluid viscosity; V is the element volume; G_i is the set of elements adjacent to the i -th element; and T_{ij} is the transmissibility between fine grids, which is defined by

$$T_{ij} = \frac{A_{ij} K_{ij}}{d_i + d_j} \quad (7)$$

where A_{ij} is the interfacial area between the grids; K_{ij} is the harmonically averaged permeability between the grids; and d is the vertical centroid-to-interface distance. q is defined by the following Peaceman equation (Peaceman, 1993):

$$q = \frac{1}{\mu} \text{WI} (p - p_{\text{wf}}) \quad (8)$$

where p_{wf} is the bottom-hole pressure (BHP) and WI is the well index, which is defined as

$$\text{WI} = \frac{2\pi K_a h}{\ln(r_e/r_w) + s} \quad (9)$$

where r_w and r_e indicate the wellbore radius and equivalent radius, respectively; h is the thickness of the connected formation; K_a is the effective permeability normal to the wellbore; s is the skin factor.

In this study, we employ closed (no-flow) outer boundaries and fixed bottom-hole pressure (BHP) well controls, implemented via the standard Peaceman model. This consistent setup establishes a unified global driving mechanism, defining the learning task as a stable permeability-to-pressure operator that aligns with the FNO's spectral inductive bias. The FVM-based loss strictly enforces these constraints, automatically satisfying zero-flux outer boundaries and embedding well controls directly into the local residuals, thereby ensuring rigorous mass conservation and physical consistency.

At the coarse scale, as shown in Fig. 1, an equivalent transmissibility is calculated for the new interface (the red portion) between any two coarse blocks after upscaling, ensuring that the flux matches the total fine-scale flux. Fig. 1 shows an example with a regular grid; in actual reservoir simulations, the fine grids and the coarse grids are typically irregular CPGs. In the coarse-scale simulation, the upscaled $\bar{T}$ is used directly, which avoids the potential deviation caused by recalculating transmissibility from Eq. (7), particularly on non-orthogonal CPGs. Therefore, the upscaled transmissibility $\bar{T}_{IJ}$ is calculated by

$$\mathcal{Q}_{IJ} = \frac{1}{\mu} \bar{T}_{IJ} (\bar{p}_I - \bar{p}_J) \quad (10)$$

where I and J are the indices of the coarse blocks; $\bar{p}_I$ and $\bar{p}_J$ are the volume-weighted average pressures of the I -th and J -th coarse grids; and $\mathcal{Q}_{IJ}$ is the flux between the coarse grids, calculated by

$$\mathcal{Q}_{IJ} = \sum_{i,j \in \Gamma_{IJ}} \frac{1}{\mu} T_{ij} (p_j - p_i) \quad (11)$$

where Γ_{IJ} is the interface between the coarse blocks. Additionally, the equivalent well index $\bar{\text{WI}}$ is calculated using

$$\bar{\text{WI}} = \frac{1}{\bar{p}_I - p_{\text{wf}}} \sum_{i \in \Omega_I} \text{WI}_i (p_i - p_{\text{wf}}) \quad (12)$$

In the numerical simulation-assisted global upscaling (NSAU) workflow, Eq. (6) is first solved at the fine scale to obtain p . The fine grids are then merged into coarse grids through re-gridding. Finally, the upscaled parameters are calculated using Eq. (10) and Eq. (12). In this study, MATLAB Reservoir Simulation Toolbox (MRST) (Lie, 2019) is used to perform both the fine-scale pressure solution and the re-gridding steps.

2.3. Neural network-assisted upscaling

In this section, we introduce NN-assisted upscaling methods, which use NNs to learn the mapping relationships in specific steps of the upscaling workflow. After sufficient training, the NN reasoning is used to replace specific computational processes. As shown in Fig. 2, the NN architecture is based on the FNO (Wang et al., 2020) and its variants for flow problems (Wen et al., 2022; Yan et al., 2025). This method utilizes the Fast Fourier Transform (FFT) (Chi et al., 2020) and suppresses high-frequency noise, thereby capturing universal low-frequency patterns among samples to achieve robust generalization. Because the global feature extraction is performed in the frequency domain, this approach is also considered to possess strong cross-resolution generalization capability (Wang et al., 2020).

The multi-layer mapping of the NN can be represented as

$$H^{(\text{in})} \leftrightarrow H^{(0)} \leftrightarrow H^{(1)} \leftrightarrow H^{(2)} \leftrightarrow \dots \leftrightarrow H^{(m)} \leftrightarrow H^{(\text{out})} \quad (13)$$

where $H^{(\text{in})}$ and $H^{(\text{out})}$ are the input and output variables of the NN, respectively; $H^{(0)}, H^{(1)}, \dots, H^{(m)}$ are the intermediate variables of the hidden layers; and m is the number of intermediate variables.

Within the framework of global transmissibility upscaling, fine-scale anisotropy is no longer equivalently calculated in the form of tensors within coarse blocks. Instead, this directional dependence is uniformly encoded as the spatial distribution and magnitude of transmissibility. In other words, the anisotropic permeability tensor can be input in a multi-channel format to the NN, but the output is always a single-channel transmissibility scalar. In this study, the input $\mathbf{K}$ is simplified to be the isotropic K to improve computational efficiency.

In Fig. 2, the first component of the NN is the input layer, with its mapping defined in Eq. (14), which embeds K and the spatial coordinates X , Y , and Z (grid-center coordinates) into a high-dimensional space:

$$H^{(0)} = W^{(\text{in})} \cdot [K, X, Y, Z] + b^{(\text{in})} \quad (14)$$

where $W^{(\text{in})}$ is the weight of the input layer and $b^{(\text{in})}$ is the bias of the input layer. The mapping equation of hidden layers is

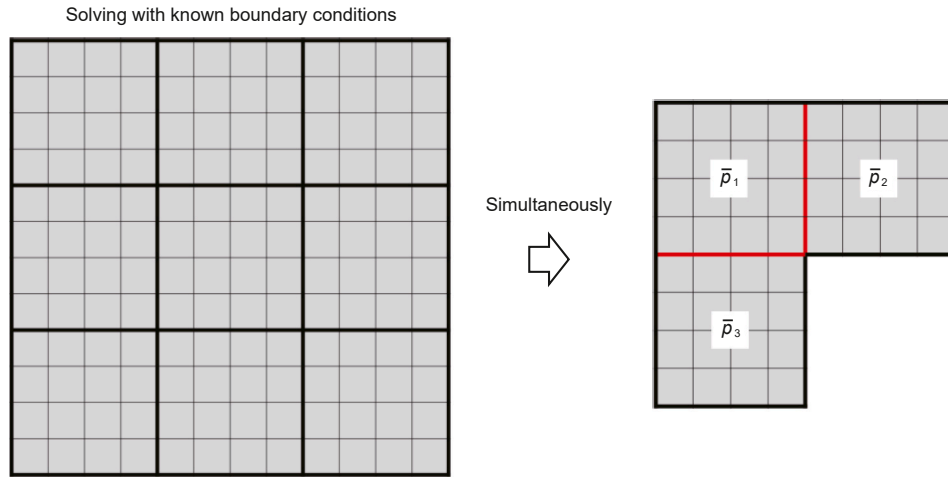


Fig. 1. Upscaling of transmissibility with known boundary condition.

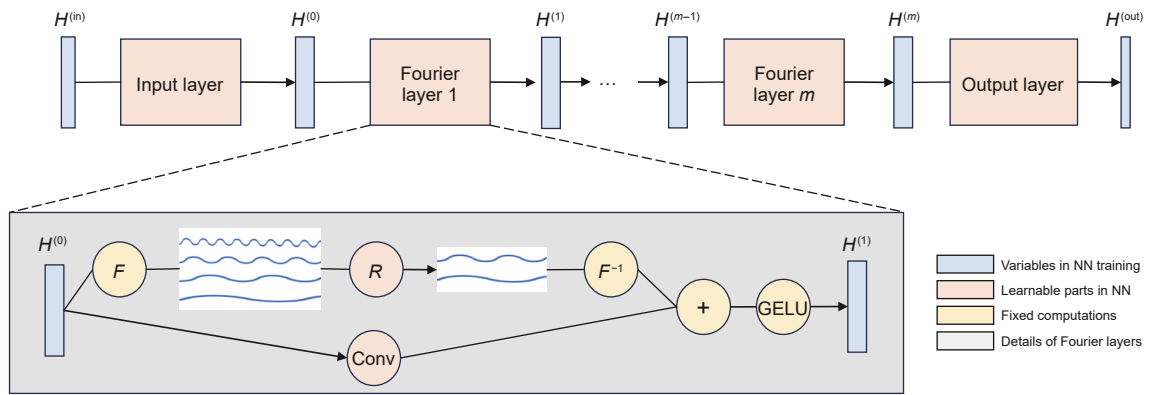


Fig. 2. Basic neural network architecture for assisting model upscaling.

$$H^{(l)} = \text{GELU}\left(\text{BN}\left(F^{-1}\left(R^{(l)} \cdot F\left(H^{(l-1)}\right)\right) + \text{Conv}^{(l)}\left(H^{(l-1)}\right)\right)\right), \forall l = 1, \dots, m-1 \quad (15)$$

where GELU is the Gaussian Error Linear Unit (Hendrycks and Gimpel, 2016), which achieves smooth activation through Gaussian distribution gating and exhibits better convergence stability and generalization capability in deep models (Devlin et al., 2019; Lee, 2023); BN is Batch Normalization; F is the Fourier Transform and F^{-1} is the inverse Fourier Transform; R is the low-pass filtering layer in the frequency domain, with its dimensions determined by the hyperparameters n_x , n_y , and n_z that control the number of modes in three spatial directions; and Conv is the convolutional layer. After multiple hidden-layer mappings, the mapping of the output layer is

$$H^{(\text{out})} = W^{(\text{out})} \cdot H^{(m)} + b^{(\text{out})} \quad (16)$$

where $W^{(\text{out})}$ denotes the output layer weights; $b^{(\text{out})}$ is the output layer bias. The output layer projects the features back into a low-dimensional space, and its can be modified to accommodate different data structures.

As shown in Fig. 3, in this study, three different NN-assisted upscaling pathways are designed to compare the effects of different mapping relationships and training strategies, including:

a data-driven permeability-to-transmissibility surrogate (K2T); a data-driven permeability-to-pressure surrogate (K2Pdata); and a physics-driven permeability-to-pressure surrogate (K2Pphy). For the data-driven methods, samples are provided by the NSAU method described in Section 2.1.

(1) Data-driven permeability-to-transmissibility surrogate

As shown in Fig. 3(b), in the K2T pathway, the NN is used to replace the mapping $K \leftrightarrow p \leftrightarrow \bar{T}$. The trained NN can be used to map a batch of fine-scale permeability fields to coarse-scale transmissibility fields Fig. 4 illustrates a 2D case; actual reservoir models are typically constructed from 3D CPGs, in which case the input permeability is expanded into a batch of 3D matrices. This pathway covers intermediate steps, forming a complete end-to-end NN prediction of the upscaling results. In addition, the new well indices are also predicted by the NN, Eq. (13) is modified as

$$[K, X, Y, Z] \leftrightarrow H^{(0)} \leftrightarrow H^{(1)} \leftrightarrow H^{(2)} \leftrightarrow \dots \leftrightarrow H^{(m)} \leftrightarrow [\hat{T}, \widehat{\text{WI}}] \quad (17)$$

where $\hat{T}$ is the NN output transmissibility; $\widehat{\text{WI}}$ is the NN output well index. In practice, $\hat{T}$ and $\widehat{\text{WI}}$ are expanded and then concatenated into a column as the NN output.

We construct the following loss function based on the mean squared error (MSE) to perform data-driven training

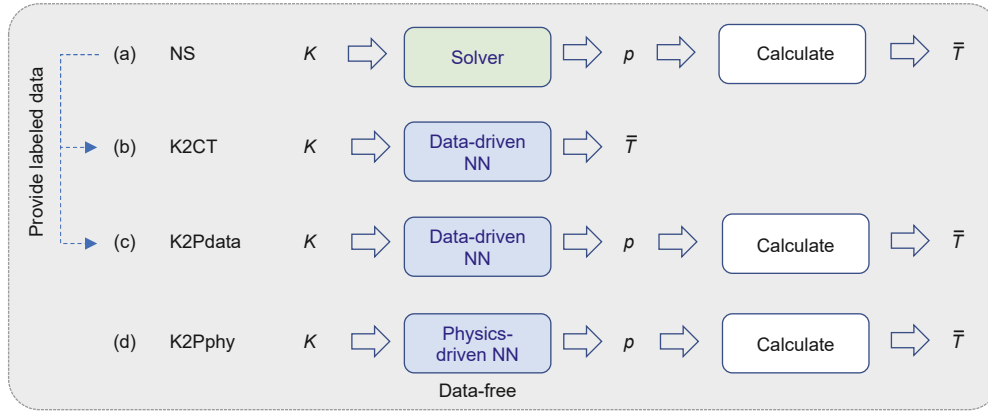


Fig. 3. Numerical simulation-assisted and neural network-assisted upscaling pathways.

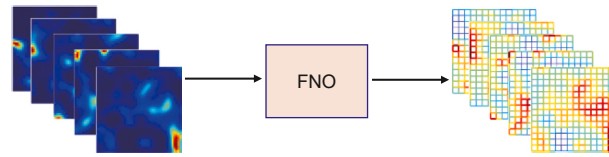


Fig. 4. Multidimensional multi-channel mapping from permeability to transmissibility.

$$L_{\text{data,K2T}} = \frac{1}{N_{\text{face}}} \sum_{l=1}^{N_{\text{face}}} (\bar{T}_l - \hat{T}_l)^2 + \frac{1}{N_{\text{well}}} \sum_{l=1}^{N_{\text{well}}} (\bar{W}_l - \hat{W}_l)^2 \quad (18)$$

where N_{face} is the number of internal faces of the coarse model; N_{well} is the total number of well segments after upscaling.

To facilitate the convergence, the following normalization is applied to scale all inputs and outputs to a range between 0 and 1. Specifically, max-min normalization is used for transmissibility, well index, and spatial coordinates, while the input permeability is logarithmically normalized using

$$\bar{K}_{\text{normal}} = \frac{\bar{K}_{\text{log}} - \bar{K}_{\text{log,min}}}{\bar{K}_{\text{log,max}} - \bar{K}_{\text{log,min}}} \quad (19)$$

where $\bar{K}_{\text{normal}}$ is the scaled permeability, $\bar{K}_{\text{log}}$ is the common logarithmic permeability; $\bar{K}_{\text{log,max}}$ and $\bar{K}_{\text{log,min}}$ are the maximum and minimum values of the common logarithmic permeability, respectively. All maximum and minimum values used for scaling are derived from the training set, and the NN outputs are subsequently denormalized using these values to obtain the actual predicted results.

(2) Data-driven permeability-to-pressure surrogate

In the K2Pdata pathway (Fig. 3(c)), the FNO is used to replace the pressure solution process $K \leftrightarrow p$, after which Eq. (10) and Eq.

(12) are employed to compute $\bar{T}$ and $\bar{W}_l$, respectively. As illustrated in Fig. 5, both the $K \leftrightarrow p$ mapping and the subsequent explicit calculation can be rapidly processed in parallel in a multi-dimensional matrix format.

The loss function of K2Pdata is defined as

$$L_{\text{data,K2Pdata}} = \frac{1}{N_{\text{grid}}} \sum_{i=1}^{N_{\text{grid}}} (p_i - \hat{p}_i)^2 \quad (20)$$

where N_{grid} is the number of fine grids; $\hat{p}$ is the pressure output by the NN.

Spatial coordinates are normalized using max-min normalization, permeability is normalized based on Eq. (19), and pressure is normalized using

$$p_{\text{normal}} = \frac{p - p_{\text{init,min}}}{p_{\text{init,max}} - p_{\text{init,min}}} \quad (21)$$

where p_{normal} is the scaled pressure; $p_{\text{init,max}}$ and $p_{\text{init,min}}$ are the maximum and minimum values of the preset pressures at the boundaries and source/sink locations, which are typically determined by observations. The outputs of the NN are subsequently denormalized based on these unified standards to obtain the actual pressure.

(3) Physics-driven permeability-to-pressure surrogate

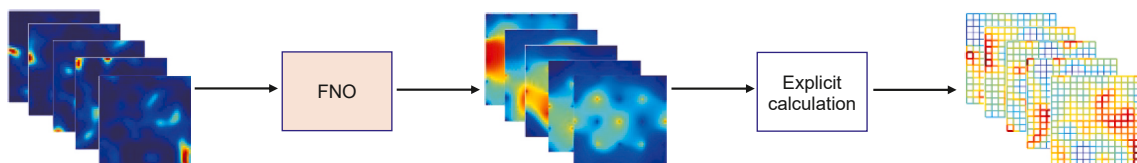


Fig. 5. Multidimensional multi-channel mapping from permeability to pressure and equivalent transmissibility calculation.

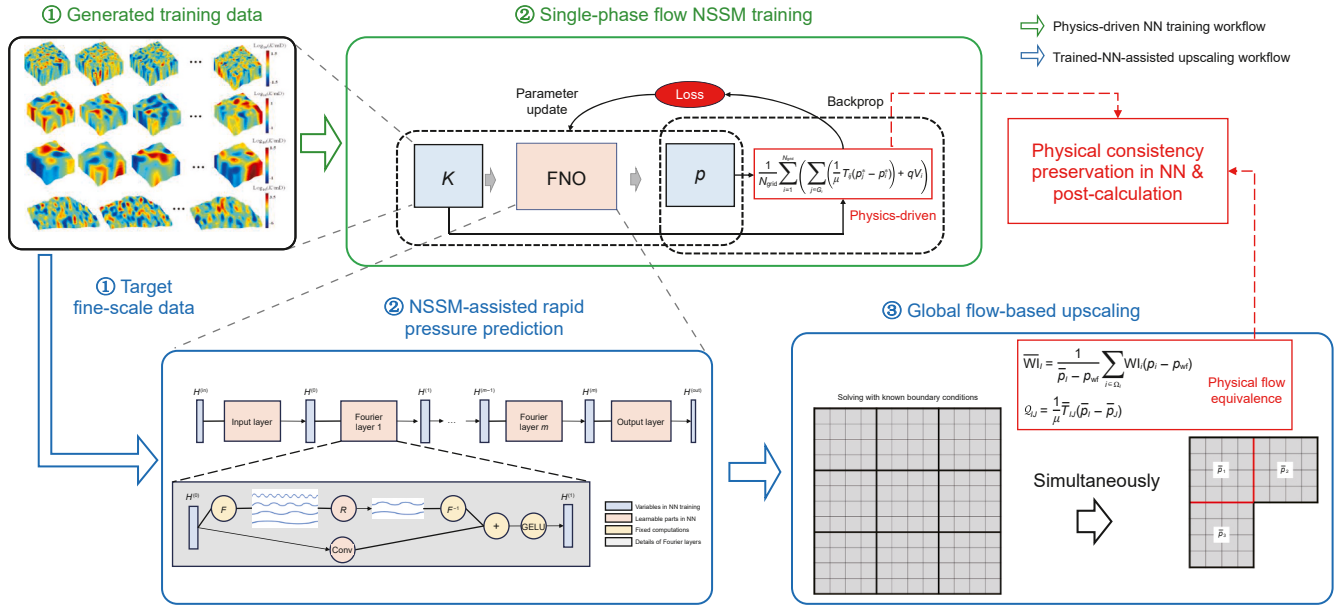


Fig. 6. Schematic diagram of the K2Pphy framework: Physics-driven training and rapid upscaling workflows.

As shown in Fig. 3(c), in the K2Pphy pathway, the step replaced by the NN is the same as in K2Pdata. The mapping result of the NN is also identical to that in Fig. 5. The difference is that the training of the FNO no longer uses data from the NSAU but instead directly performs data-free, physics-informed training, as shown in Fig. 6. In the PDE-based learning, the physics-constrained loss function (Raissi et al., 2019; Cai et al., 2021; Donnelly et al., 2024) is typically defined as

$$L(\theta) = L_{\text{res}}(\theta) + L_{\text{ic}}(\theta) + L_{\text{bc}}(\theta) \quad (22)$$

where θ is the NN parameters; L_{res} is the structural loss of the PDE residual; L_{ic} and L_{bc} are the initial condition loss and boundary condition loss, respectively.

In this study, the NN is established to surrogate the pressure solution in steady-state single-phase flow, thus $L_{\text{ic}}(\theta)$ does not exist. Moreover, the zero-flux Neumann boundary condition is automatically satisfied under FVM discretization by Eq. (6). Consequently, the physical loss function used for NN training is defined as

$$L_{\text{res}, \text{K2Pphy}} = \frac{1}{N_{\text{grid}}} \sum_{i=1}^{N_{\text{grid}}} \left(\sum_{j \in G_i} \left(\frac{1}{\mu} T_{ij} (\hat{p}_j - \hat{p}_i) \right) + qV_i \right) \quad (23)$$

The inputs and outputs of K2Pphy are normalized and denormalized in the same manner as those of K2Pdata.

In the training of K2T, K2Pdata, and K2Pphy, mini-batch gradient descent (Khairat et al., 2017) is employed. In each gradient descent iteration, a sub-batch of samples is randomly selected for mapping and loss computation, followed by one backpropagation and parameter update. When all samples have been processed through, one epoch is completed. During the training process, the model cycles through multiple epochs until the convergence criterion is satisfied. This stochastic process is considered to enhance the generalization capability of the model. Notably, as illustrated in Fig. 6, physical consistency is enforced throughout the entire K2Pphy workflow, encompassing training, inference, and upscaling. This holistic physics-informed regularization serves as a critical mechanism to further enhance the model's robustness against unseen geological scenarios.

2.4. The model settings and evaluation methods

To validate the capability of the methodology, two CPG models are provided. The first model is the synthetic model (Fig. 7(a)), which covers an area of approximately 675 m × 675 m × 60 m. In the horizontal direction, 46 × 46 pillars are established (corresponding to 45 × 45 planar elements), and the model is subdivided vertically into 6 layers, resulting in a total of 12,150 active cells. The horizons are generated through interpolation, which influenced by stratum undulations, results in variations in cell thickness and surface dip angle with spatial position. The wells in the synthetic model follow a five-spot well pattern, including injection wells (sources) I1, I2, ..., I7 and production wells (sinks) P1, P2, ..., P18. In re-gridding process, the fine grids are partitioned into a logical 15 × 15 × 3, followed by grid merging. The second model is the standard SAIGUP model (Howell et al., 2008), as shown in Fig. 7(b). It covers an area of approximately 2925 m × 9000 m × 593 m and contains multiple faults. In the horizontal direction, 41 × 121 pillars are constructed, and the vertical direction is subdivided into 20 layers, for a total of 78,720 active cells. The wells in the SAIGUP model follow a line-drive pattern, including sources I1, I2, ..., I26 and sinks P1, P2, ..., P22. The CPGs of the SAIGUP model is more distorted, and the cell sizes vary more significantly. For structurally complex practical reservoir models, vertical merging usually requires specialized analysis and decision-making based on geological and production data, including reasonable stratigraphic division and development-layer merging according to geological structure, sedimentary facies distribution, hydrocarbon occurrence, and vertical heterogeneity (Krishnamurthy et al., 2024; Yi et al., 2022; Fu et al., 2024). Therefore, for the SAIGUP model, only planar grid coarsening is applied, with the logical space partition set to 13 × 40 × 20. The above re-gridding is carried out using the upscaling module in the MRST (Lie, 2019). This module includes algorithms to identify and partition grids across faults to approximate the original geological structure as closely as possible, as well as automatic selection algorithms to merge fine grids as uniformly as possible, thereby improving the convergence of subsequent simulations. The boundaries of both CPG models are set as closed boundaries with zero flux. The initial reservoir

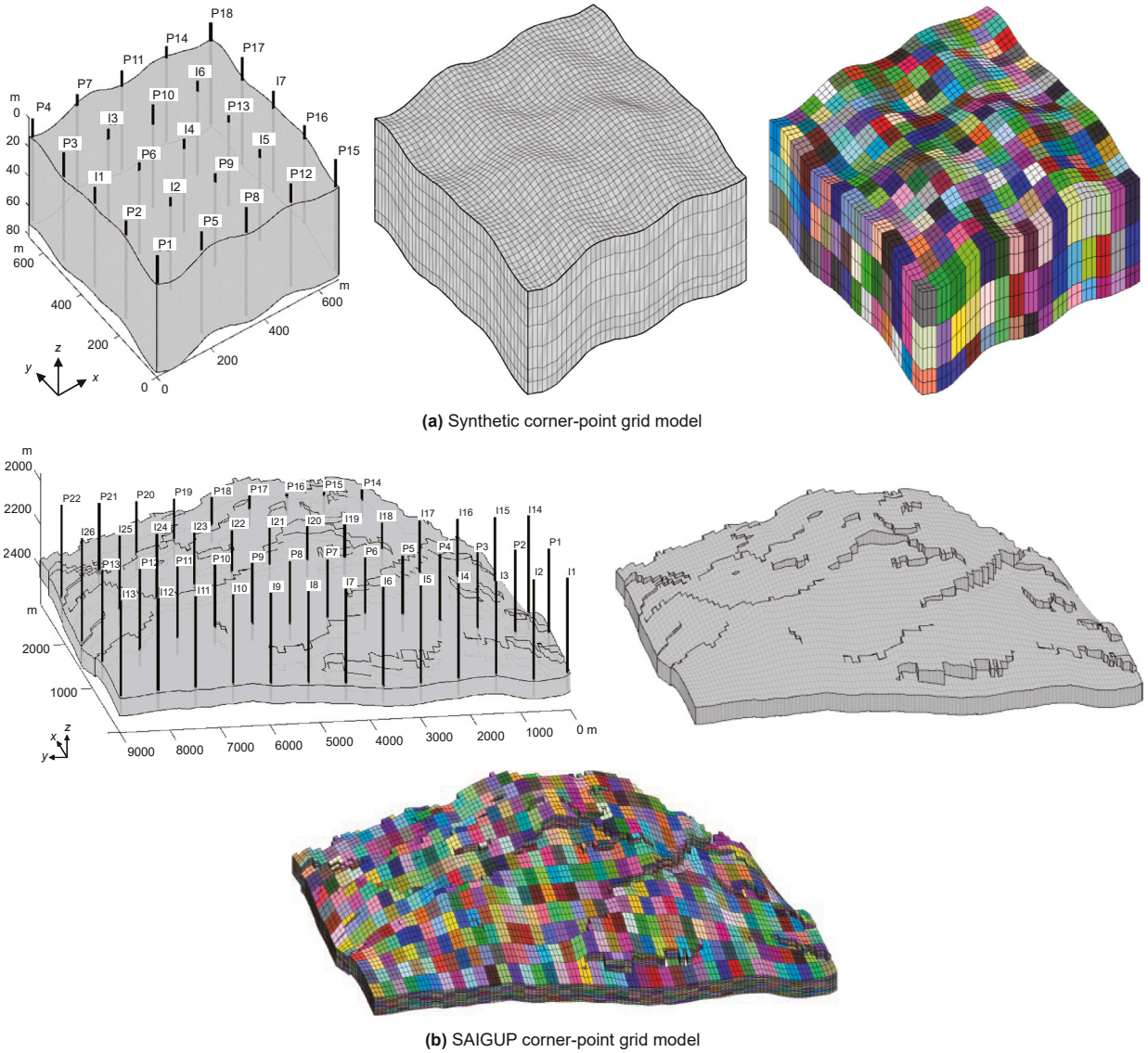


Fig. 7. Well configuration, fine-grid, and coarse-scale re-gridding of numerical models.

pressure is set to 30 MPa. All injection and production wells are perforated across the entire formation and are controlled by BHP, with injectors at 20 MPa and producers at 40 MPa. As shown in Fig. 7, the well patterns are distributed across the entire block, which drives global flow while also making more distinct local flow features. The porosity of both models is uniformly set to 0.1, and the fluid viscosity is set to 1 mP·s.

We sample random heterogeneous permeability fields that follow a log-normal distribution using the Gaussian random field (GRF) method. For N_{grid} sampling points in the spatial domain $D \subset \mathbb{R}^3$, the log-permeability field is defined as

$$K_{\log}(i) \sim \mathcal{GP}(\mu, C(\chi_i, \chi_j)) \quad (24)$$

where $K_{\log}$ denotes the common logarithmic permeability field of the fine grid, χ represents the spatial location, and $\mathcal{GP}$ denotes a Gaussian random process with mean μ . The kernel function C is defined as

$$C(\chi_i, \chi_j) = \sigma^2 \exp\left(-\left(\frac{r_{ij}}{\ell}\right)^2\right) + \tau^2 \delta_{ij} \quad (25)$$

where σ^2 is the variance; $r_{ij} = \|\chi_i - \chi_j\|$ represents the Euclidean distance; ℓ is the correlation length; τ^2 is the nugget; δ_{ij} is the Kronecker delta. The semivariogram γ is defined as

$$\gamma(r) = \sigma^2 \left(1 - \exp\left(-\left(r/\ell\right)^2\right)\right) \quad (26)$$

After setting the GRF model parameters, the common logarithmic permeability field can be sampled from the distribution

$$K_{\log} \sim \mathcal{N}(\mu \mathbf{1}, \mathbf{C}) \quad (27)$$

where $\mathbf{C}$ is the covariance matrix, $\mathbf{C} \in \mathbb{R}^{N_{\text{grid}} \times N_{\text{grid}}}$, $\mathbf{C}_{ij} = C(\chi_i, \chi_j)$; $\mathbf{1}$ is the all-ones vector. Each realization is sampled as

$$K_{\log} = \mu \mathbf{1} + \mathbf{L}\xi \quad (28)$$

where $\mathbf{L}$ is the lower triangular matrix obtained from the Cholesky decomposition, $\mathbf{C} = \mathbf{L}\mathbf{L}^T$; ξ is a randomly generated vector with $\xi \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$; $\mathbf{I}$ is the identity matrix. By repeatedly generating ξ , permeability realizations $K^{(1)}, \dots, K^{(N_{\text{real}})}$ are obtained, where N_{real} is the number of samples. Finally, the permeability field for simulation is obtained through $K = 10^{K_{\log}}$. All sampling procedures in this study are performed using Geostatistical Tools (GSTools) (Müller et al., 2022).

As shown in Table 1, five cases are designed to compare various upscaling pathways from different perspectives to illustrate the applicability in reservoir model upscaling. Multiple numerical models are employed, including Sy1, Sy2, Sy3, Sy4, and Sy5 based on the synthetic CPG model, as well as SA1 based on SAIGUP. In permeability sampling, the mean of $K_{\log}$ is set to 0 and the nugget is set to 0 for all cases. More detailed parameter settings and statistical results are provided in Table 2. The spatial distribution of $K_{\log}$ for some of the realizations are shown in Fig. 8.

Case 1 is designed on the basic Sy1 model to compare the upscaling performance of NSAU, K2T, K2Pdata, and K2Pphy. 1000 samples are generated, with 500 randomly selected as the training set and the remaining samples as the test set. At the same time, Case 1 also serves as the baseline for comparison with Case 2 and Case 3, which are used to test the generalization capabilities in

heterogeneous unknown parameter spaces. Subcases Case 2.1 and Case 2.2 use the pretrained NN to inference towards lower (Sy2) and higher (Sy3) permeability variance, respectively (a larger variance usually indicates greater heterogeneity (Rubin, 2003)). Case 3.1 and Case 3.2 inference towards smaller-scale (Sy4) and larger-scale (Sy5) features, respectively. Larger-scale features usually imply the presence of more continuous geological structures, such as fluvial facies or deep-sea sediments, whereas smaller-scale features correspond to environments with stronger local variations, such as alluvial fan deposits. It should be noted that Case 2 and Case 3 do not involve a training process (N/A in Table 2 indicates no training set). Therefore, 500 samples are generated as the test set. We load the K2T, K2Pdata, and K2Pphy models trained in Case 1, which are set to evaluation mode and used to inference on the test set, in order to assess their extrapolation generalization ability. In other words, Case 1, Case 2, and Case 3 can be viewed as a unified whole, comprising one training set (500 samples randomly selected from Sy1 in Case 1) and five test sets. These test sets include: the remaining 500 samples from Sy1 in Case 1 (to test basic, in-distribution inference capability), and the 500 samples from each of Sy2 and Sy3 in Case 2, as well as Sy4 and Sy5 in Case 3 (to test advanced, out-of-distribution inference capability). The differences in parameter settings for generating all training and test

Table 1
Parameter settings for all cases.

Grid	Case	Upscaling method	Dataset			Post evaluation indicator
			Numerical model	Log permeability variance	Correlation range, m	
Synthetic	1	NSAU, K2T, K2Pdata, K2Pphy	Sy 1	0.3	75	$Q_{1\text{phase}}$
	2.1	NSAU, K2T, K2Pdata, K2Pphy	Sy 2	0.03	75	$Q_{1\text{phase}}$
	2.2		Sy 3	3		
	3.1	NSAU, K2T, K2Pdata, K2Pphy	Sy 4	0.3	30	$Q_{1\text{phase}}$
	3.2		Sy 5		120	
	4	NSAU, K2T, K2Pdata, K2Pphy	Sy 1	0.3	75	$Q_{\text{inj}}, Q_{\text{pro}}, f_w$
SAIGUP	5	NSAU, K2Pphy	SA 1	0.3	300	$Q_{1\text{phase}}, Q_{\text{inj}}, Q_{\text{pro}}, f_w$

Table 2
Numerical model parameter settings.

Grid	Numerical model	Log permeability variance	Correlation range, m	All fields (min–max), mD	Field average (min–max), mD	Samples (train, test)
Synthetic	Sy1	0.3	75	0.0028–1281.68	0.77–10.95	500, 500
	Sy2	0.03	75	0.15–9.61	0.83–1.37	N/A, 500
	Sy3	3	75	8.68e–9–4.49e+6	8.94–8.35e+6	N/A, 500
	Sy4	0.3	30	0.0025–4.01	1.94–2.64	N/A, 500
	Sy5	0.3	120	0.031–52.77	1.73–2.61	N/A, 500
	SA1	0.3	300	0.0031–399.37	1.34–4.37	500, 500

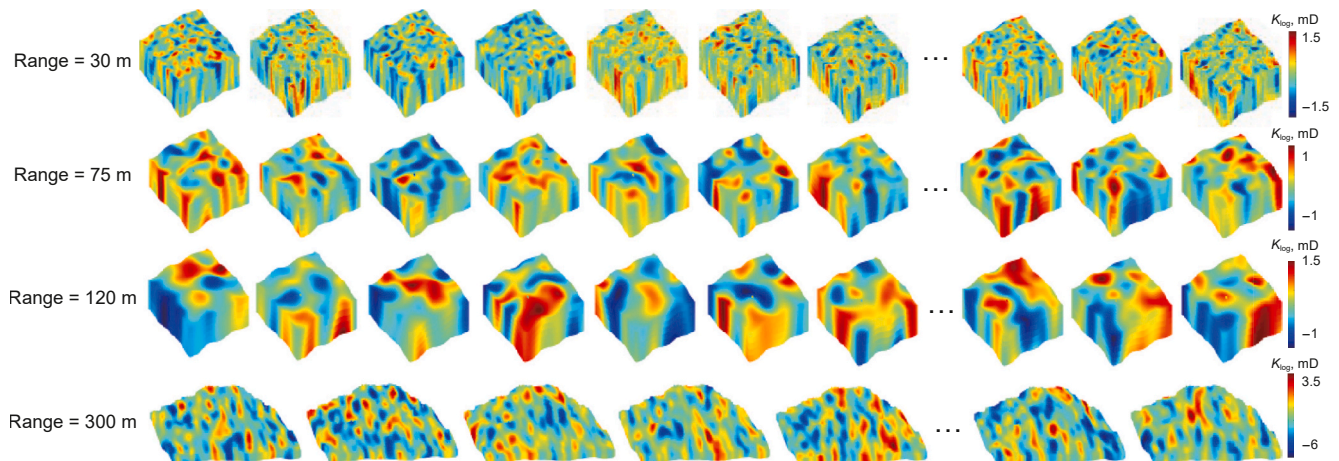


Fig. 8. Sampled common logarithmic permeability ($K_{\log}$) realizations.

set samples are shown in the “Log permeability variance” and “Correlation range” columns of Table 2. Since the generation process involves randomness, the “All Fields range” and “Field average” of the generated results are also summarized in the table to visually inspect the differences among samples.

In the comparison of Case 1, Case 2, and Case 3, the upscaling performance is evaluated through unsteady single-phase slightly compressible flow simulations. The detailed numerical model settings are presented in Appendix A. We employ the cumulative liquid production $Q_{1\text{phase}}$ (a well-related dynamic metric) to record the cumulative production history of all well segments, which is also a key indicator for uncertainty assessment in practical reservoir simulations (Floris et al., 2001; Cao et al., 2016). $Q_{1\text{phase}}$ can naturally accommodate different grid resolutions, thereby enabling intuitive quantitative comparison between upscaling results and the fine-scale accurate solutions. A higher degree of agreement in $Q_{1\text{phase}}$ indicates the better approximation of the fine-scale flow characteristics, ensuring more reliable evaluation based on the upscaled model. The calculation equation of $Q_{1\text{phase}}$ is

$$Q_{1\text{phase}}(N_{\text{time}}) = \sum_{\tau=1}^{N_{\text{time}}} \sum_{n=1}^{N_{\text{well}}} q_{n,\tau} \Delta t_{\tau} \quad (29)$$

where $q_{n,\tau}$ is the production rate of the n -th well segment under reservoir conditions at time step τ ; t is time; Δt_{τ} is the step size of the τ -th time step; N_{time} is the number of time steps.

In Case 4, the upscaling performance is further evaluated through two-phase slightly compressible flow simulations. We aim to intuitively qualify the upscaling results with more well-related metrics (Li et al., 2021; Tegnander and Gimse, 1998), including the cumulative water injection Q_{inj} , the cumulative

liquid production Q_{pro} , and the water cut f_w . The detailed model settings of the two-phase flow are presented in Appendix B. The calculation equation of Q_{inj} is

$$Q_{\text{inj}}(N_{\text{time}}) = \sum_{\tau=1}^{N_{\text{time}}} \sum_{n=1}^{N_{\text{inj,well}}} q_{\text{inj},n,\tau} \Delta t_{\tau} \quad (30)$$

where $q_{\text{inj},n,\tau}$ is the injection rate of the n -th injection well segment under reservoir conditions at τ -th time step; $N_{\text{inj,well}}$ is the total number of injection well segments. The calculation formula of Q_{pro} is

$$Q_{\text{pro}}(N_{\text{time}}) = \sum_{\tau=1}^{N_{\text{time}}} \sum_{n=1}^{N_{\text{pro,well}}} q_{\text{pro},n,\tau} \Delta t_{\tau} \quad (31)$$

where $q_{\text{pro},n,\tau}$ is the production rate of the n -th production well segment under reservoir conditions at τ -th time step; $N_{\text{pro,well}}$ is the total number of production well segments. The calculation formula of f_w is

$$f_w(\tau) = \frac{\sum_{n=1}^{N_{\text{pro,well}}} q_{w,n,\tau}}{\sum_{n=1}^{N_{\text{pro,well}}} q_{\text{pro},n,\tau}} \quad (32)$$

where $q_{w,n,\tau}$ is the water production rate of the n -th production well segment under reservoir conditions at τ -th time step.

Finally, in Case 5, we compare the accuracy of the NSAU and K2Pphy on the authentic SA1 model, along with the computational efficiency of their upscaling workflows, with the aim of evaluating the adaptability to practical problems. It should be noted that in Case 5, the data-driven K2T and K2Pdata are not further compared.

Table 3
Neural network parameter settings.

Grid	NN	Structure				Training		
		Input layer	Fourier layers	Output layer	Parameters	Driving method	Optimization strategy	Resource usage
Synthetic	K2T	Input dim: $4 \times 45 \times 45 \times 6$ Output dim: $10 \times 45 \times 45 \times 6$ (Linear layer)	Input dim: $10 \times 45 \times 45 \times 6$ Output dim: $10 \times 45 \times 45 \times 6$ Number of Fourier layers: 3 Mode (x,y,z): 10,10,2	Input dim: $10 \times 45 \times 45 \times 6$ Output dim: 1×1785 (Flatten & Linear layer)	21,929,986	Data-driven	Optimizer: Adam Learning rate: 0.001 Mini-batch size: 4 Shuffle: True Epochs: 1000	Memory: 1629 MiB Utilization: 50%–60%
	K2Pdata	Input dim: $4 \times 45 \times 45 \times 6$ Output dim: $10 \times 45 \times 45 \times 6$ (Linear layer)	Input dim: $10 \times 45 \times 45 \times 6$ Output dim: $10 \times 45 \times 45 \times 6$ Number of Fourier layers: 3 Mode (x,y,z): 10,10,2	Input dim: $10 \times 45 \times 45 \times 6$ Output dim: $1 \times 45 \times 45 \times 6$ (Linear layer)	240,451	Data-driven		Memory: 1209 MiB Utilization: 40%–50%
	K2Pphy	Input dim: $4 \times 45 \times 45 \times 6$ Output dim: $10 \times 45 \times 45 \times 6$ (Linear layer)	Input dim: $10 \times 45 \times 45 \times 6$ Output dim: $10 \times 45 \times 45 \times 6$ Number of Fourier layers: 3 Mode (x,y,z): 10,10,2	Input dim: $10 \times 45 \times 45 \times 6$ Output dim: $1 \times 45 \times 45 \times 6$ (Linear layer)	240,451	Physics-driven		Memory: 1306 MiB Utilization: 40%–50%
SAIGUP	K2Pphy	Input dim: $4 \times 40 \times 120 \times 20$ Output dim: $10 \times 40 \times 120 \times 20$ (Linear layer)	Input dim: $4 \times 40 \times 120 \times 20$ Output dim: $4 \times 40 \times 120 \times 20$ Number of Fourier layers: 3 Mode (x,y,z): 10,10,4	Input dim: $4 \times 40 \times 120 \times 20$ Output dim: $1 \times 40 \times 120 \times 20$ (Linear layer)	480,451	Physics-driven		Memory: 3993 MiB Utilization: 70%–80%

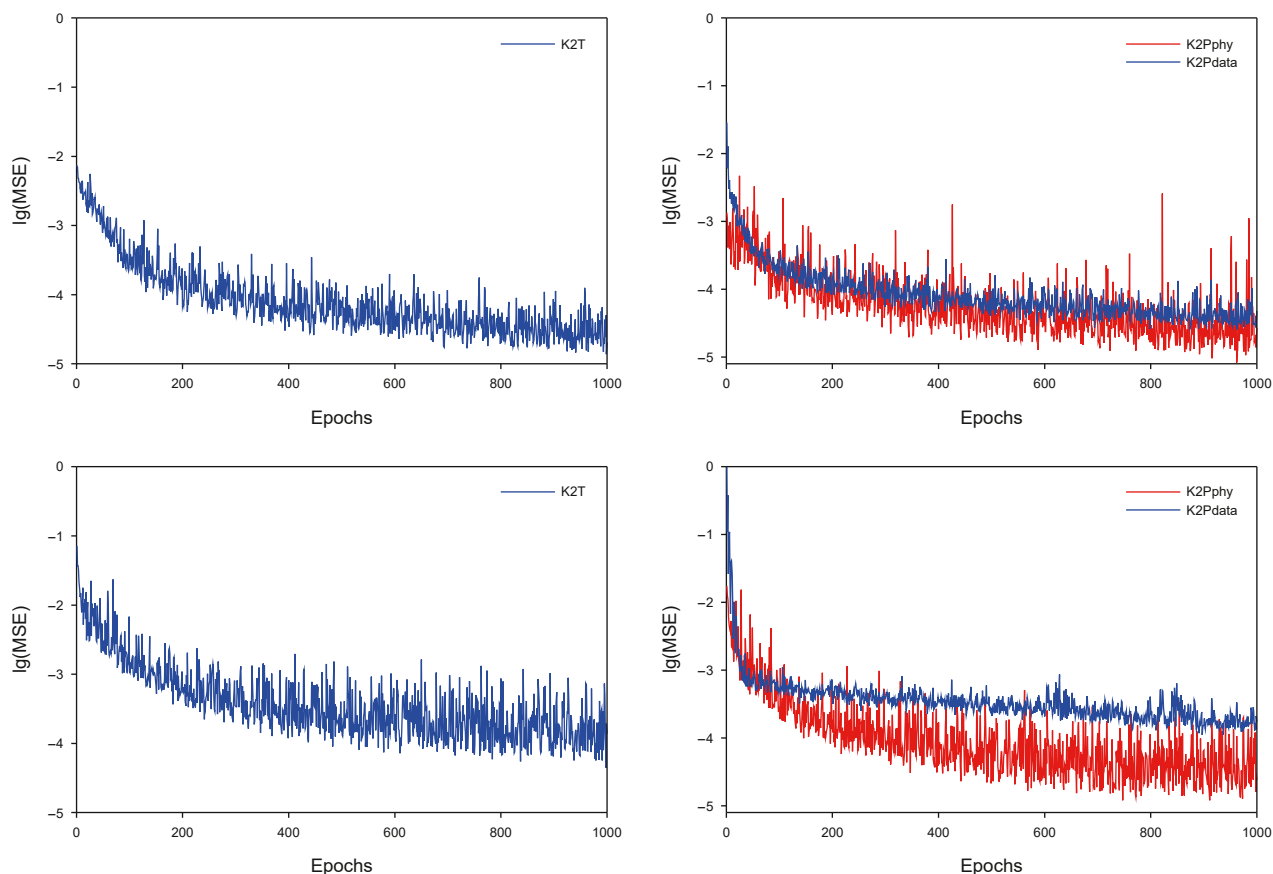


Fig. 9. Comparison of training loss curves for different methods under the synthetic CPG model (up: train; down: test).

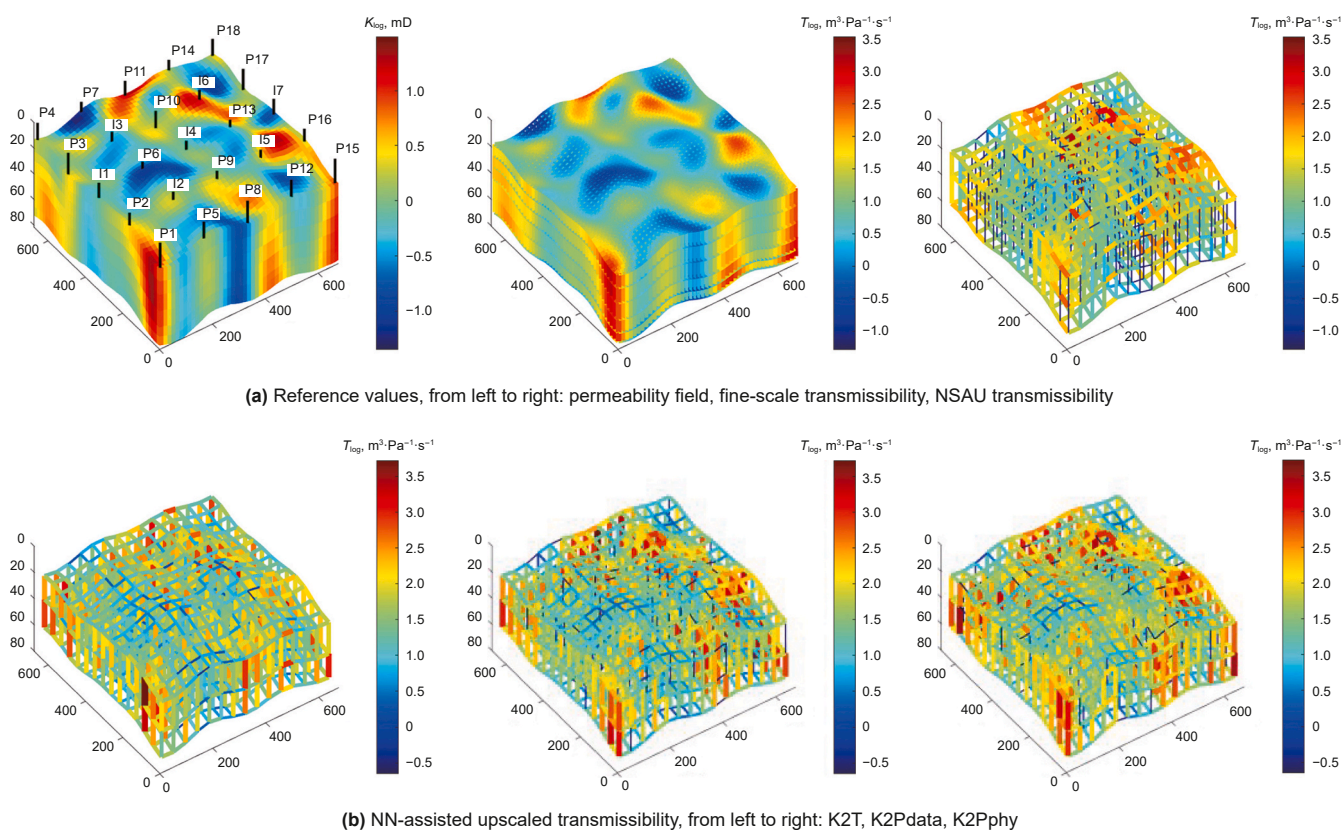


Fig. 10. Comparison of transmissibility spatial distributions of a test sample under the synthetic CPG model. $T_{\log}$ presents the logarithmic permeability.

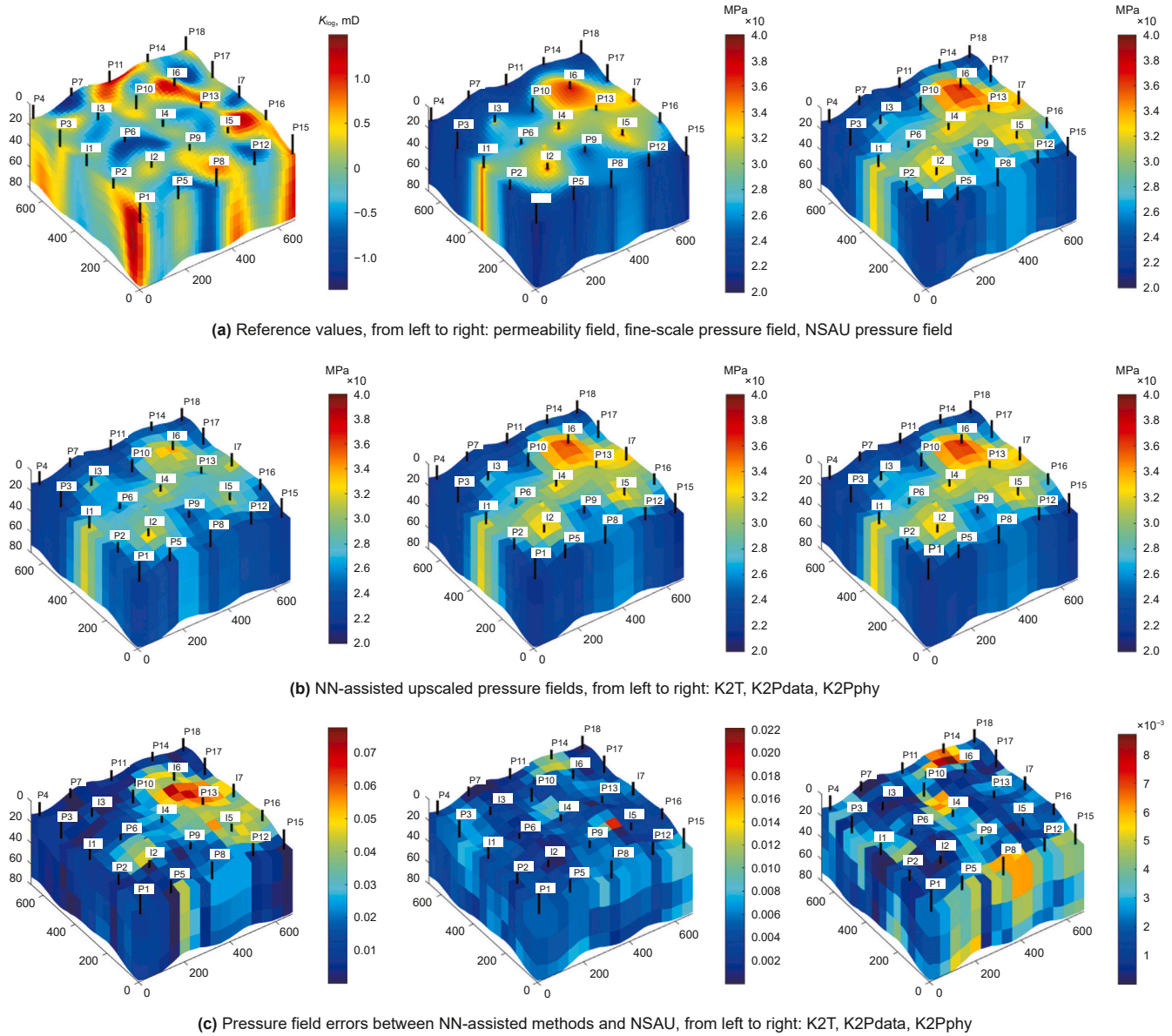


Fig. 11. Comparison of single-phase steady-state pressure field distributions under the synthetic CPG model.

The most obvious reason is their reliance on simulation data prevents them from offering any efficiency advantage, while the discussion on accuracy has already been sufficiently conducted in Case 1 to Case 4.

In this study, all NN models are constructed based on the PyTorch (version 1.12.1) framework (Ketkar and Moolayil, 2021). As shown in Table 3, the dimensions of the input and output layers of each NN are matched to the spatial dimensions of the numerical model, and these two layers perform the operations of dimensional upscaling and downscaling, respectively. The hidden layers of each NN consist of a cascaded structure of three Fourier layers, with each layer performing nonlinear mapping in high-dimensional space to approximate complex analytical relationships. For the synthetic model, the first 10 low-frequency modes are retained in the x and y directions, and the first 2 low-frequency modes are retained in the z direction; for the SAIGUP model, the first 10 low-frequency modes are retained in the x and y directions, and the first 4 low-frequency modes are retained in the z direction. All NN models are initialized using Xavier method (Glorot and Bengio, 2010) and trained with Adam (Kingma and Ba, 2014) for

1000 epochs. No additional learning rate decay or early stopping strategies are employed during training. Overfitting is structurally mitigated by the FNO's low-rank spectral truncation, which acts as implicit regularization against high-frequency noise. Furthermore, the operator-learning formulation inherently favors smooth, physically plausible solutions. Empirically, the consistent error convergence observed on the strictly isolated test set confirms that no overfitting occurred during the training process. In addition, to ensure physical consistency, our method integrates a dual mechanism: The FVM-based loss strictly enforces discrete mass conservation at the control-volume level, while the FNO architecture structurally aligns with the global nature of elliptic operators. This combination ensures the model rigorously solves the governing equations (with residuals $< 1e-4$) rather than merely fitting empirical data. All cases adopt a single general hyperparameter configuration manually tuned and presented in Table 3. We explicitly present a sensitivity analysis of the synthetic model in Appendix C to validate these settings (consistent trends were also observed for the SAIGUP model). We also test several hyperparameter optimization methods (Kingma and Ba, 2014; Akiba

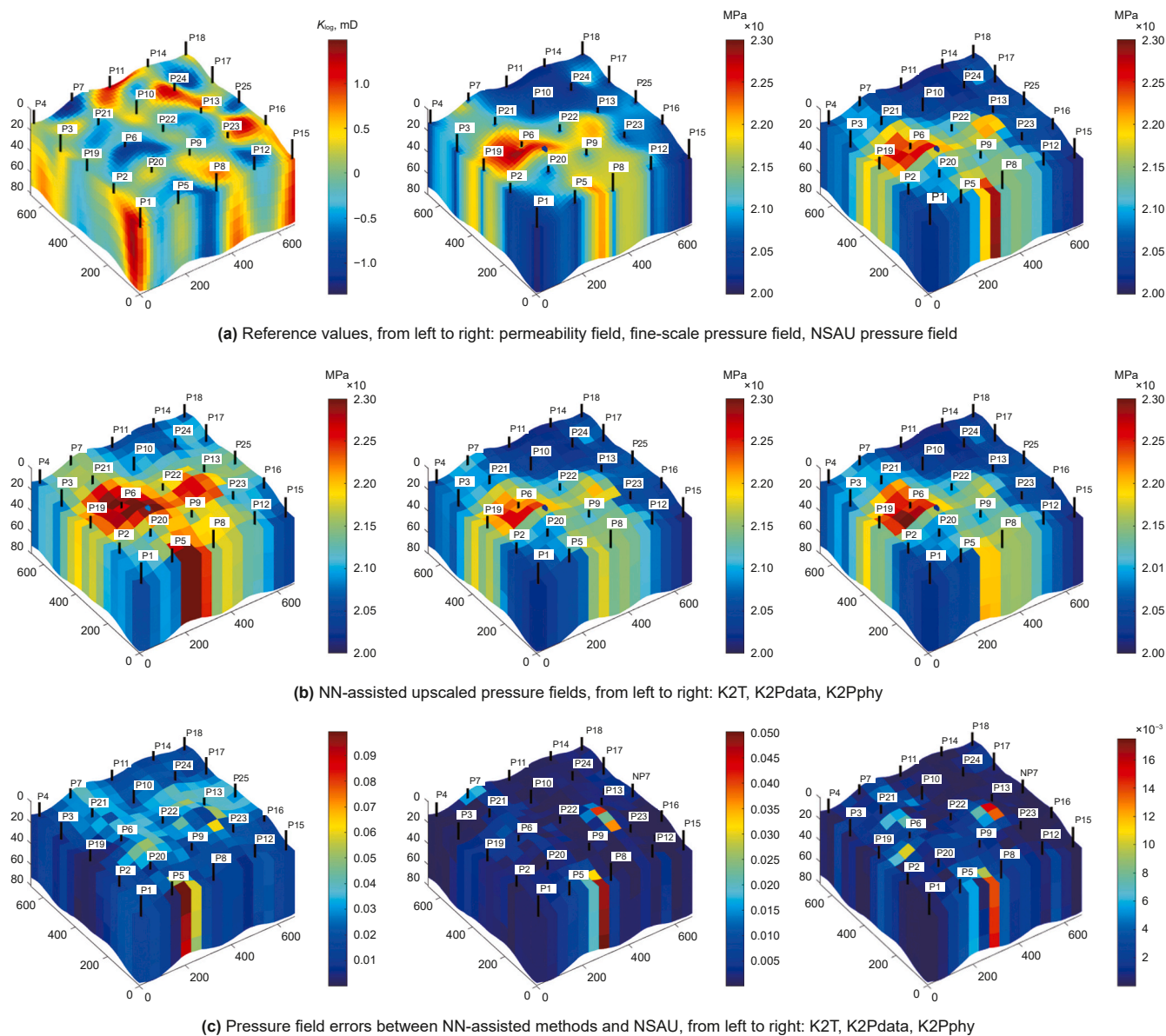


Fig. 12. Comparison of single-phase unsteady-state pressure distributions under the synthetic CPG model.

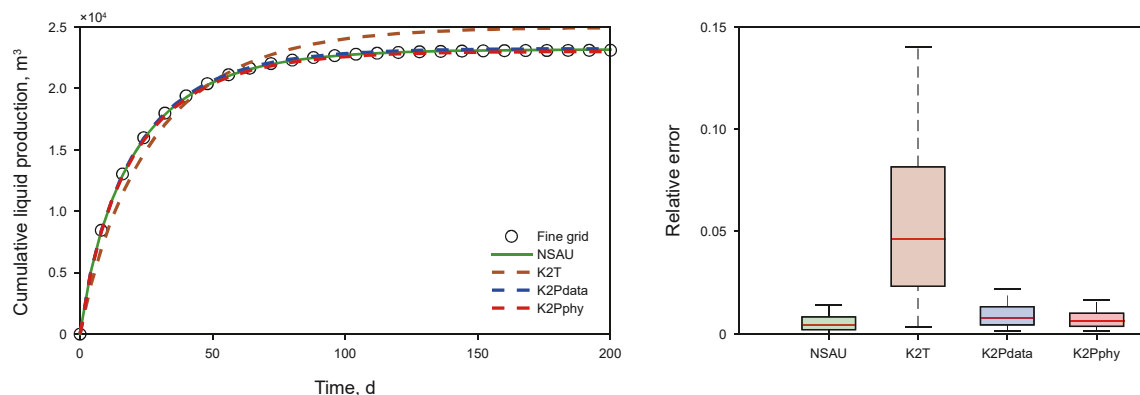


Fig. 13. Single-phase unsteady cumulative production curves and error distributions under the synthetic CPG model.

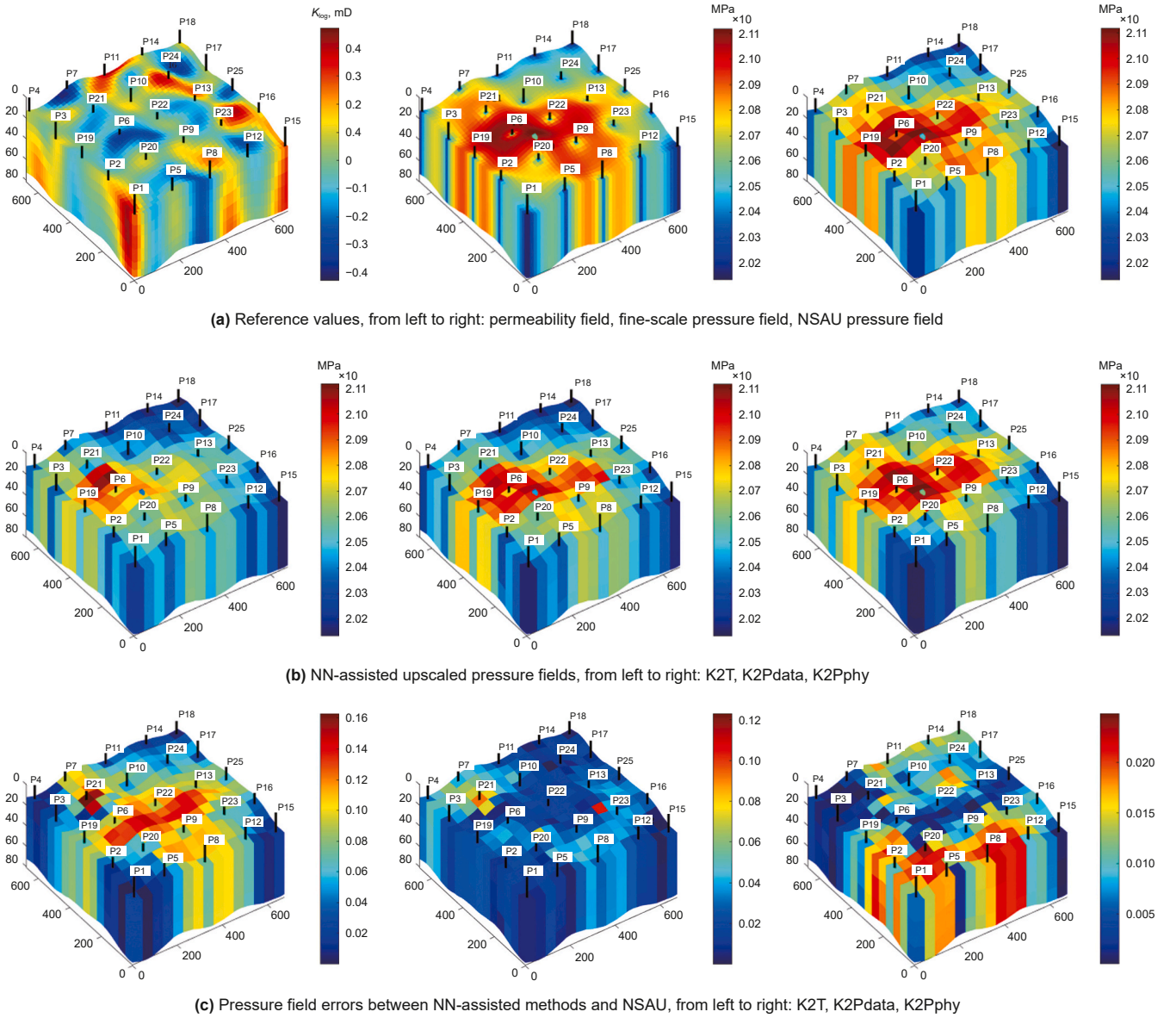


Fig. 14. Single-phase unsteady pressure field distributions of inference toward lower permeability variance under the synthetic CPG model.

et al., 2019), which can provide some minor improvements, but do not alter our main conclusions.

In the analysis of experimental results, for the quantitative indicators that can be directly compared with the fine-scale model ($Q_{1\text{phase}}$, Q_{inj} , Q_{pro} , f_w), box plots or histograms of the test-set relative error (relative to the simulated value at the final time step) are plotted. Pressure fields and saturation fields are also plotted to visually analysis the spatial distribution consistency. In addition, the flow field error between the NN-based results and the NSAU results is calculated with the following equation:

$$D_{\text{error},I,\tau} = \frac{|D_{\text{NN},I,\tau} - D_{\text{NSAU},I,\tau}|}{D_{\text{NSAU},\text{max},\tau}} \quad (33)$$

where D is the flow field, which can be either the saturation field or the pressure field; D_{NN} is the simulated result of the NN-assisted model; D_{NSAU} is the simulated result of the NSAU model; $D_{\text{NSAU},\text{max},\tau}$ is the maximum value of the NSAU model at τ -th time step.

3. Results and discussions

In this section, we first compare the basic performance of four upscaling methods (NSAU, K2T, K2Pdata, and K2Pphy) on the synthetic model with moderate heterogeneity and medium-scale features. Then, we design two sets of heterogeneous extrapolation experiments to test the generalization capabilities of different NN-assisted methods, which represent a higher-level generalization ability test. Next, we design comparisons of the simulation accuracy of different upscaling results in the more complex two-phase flow problems. Finally, we test the applicability of the new method on more complex CPGs of the authentic SAIGUP model. It should be noted that all loss curves presented in this section belong to the training set, used to observe whether the loss continues to decrease during training and the magnitude of oscillations. To enable comparison on the same scale, the loss values of all K2P models are converted into transmissibility loss. Specifically, while the NN outputs pressure, a loss is simultaneously calculated in real time based on Eq. (18), which is only used to monitor the

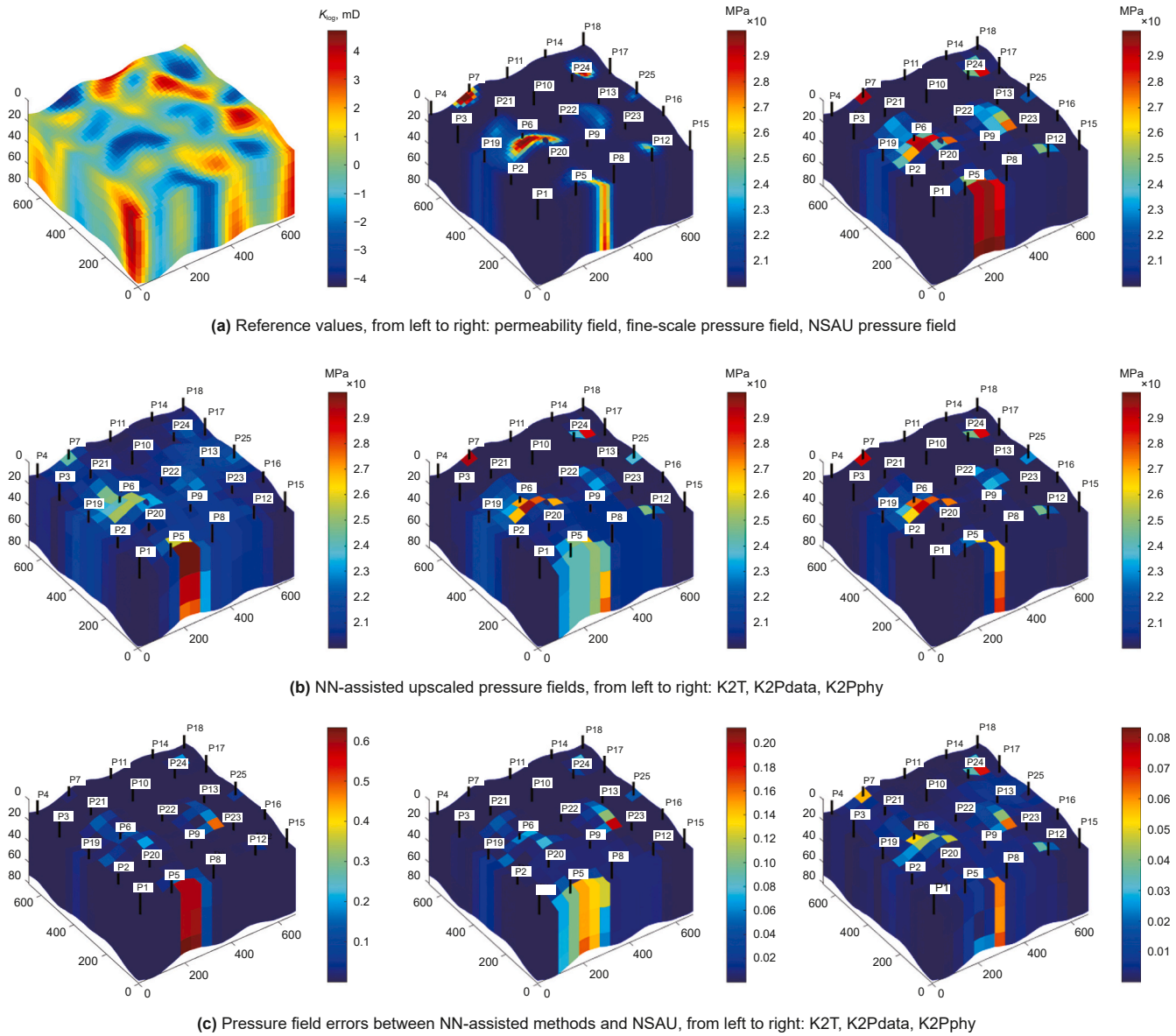


Fig. 15. Single-phase unsteady pressure field distributions toward higher permeability variance under the synthetic CPG model.

Table 4

Upscaling error of inference under different permeability variances in the test set of the synthetic CPG model.

Metrics	Cases	NSAU	K2T	K2Pdata	K2Pphy
Average pressure field errors (D_{error})	Toward lower variance	N/A	0.0647	0.0493	0.0117
	In-distribution inference	N/A	0.0387	0.0191	0.0072
	Toward higher variance	N/A	0.2524	0.0864	0.0342
Mean value of final step relative error of $Q_{1\text{phase}}$	Toward lower variance	0.0028	0.0258	0.0098	0.0051
	In-distribution inference	0.0059	0.0589	0.0113	0.0086
	Toward higher variance	0.0235	0.1925	0.1053	0.0299

training status, whereas the backpropagation and parameter updates of the K2Pdata and K2Pphy models are still based on Eq. (20) and Eq. (23), respectively. All other figures and tables in this section belong to the test set, including all physical fields (permeability, transmissibility, pressure, and saturation; all time-dependent fields plot their final time step results), production curves ($Q_{1\text{phase}}$, Q_{inj} , Q_{pro} , and f_w), error statistics plots (box plots and histograms), as well as error statistics tables.

3.1. Synthetic model training and in-distribution inference

As shown in Fig. 9, during the training of the NNs, the loss exhibits a continuous decreasing trend. Both the K2T and K2Pphy methods achieve a loss close to 1.0×10^{-5} , which is lower than that of the K2Pdata method. This indicates that the numerical errors of the former two methods relative to the NSAU method are smaller. Throughout the training process, the K2Pphy method shows the

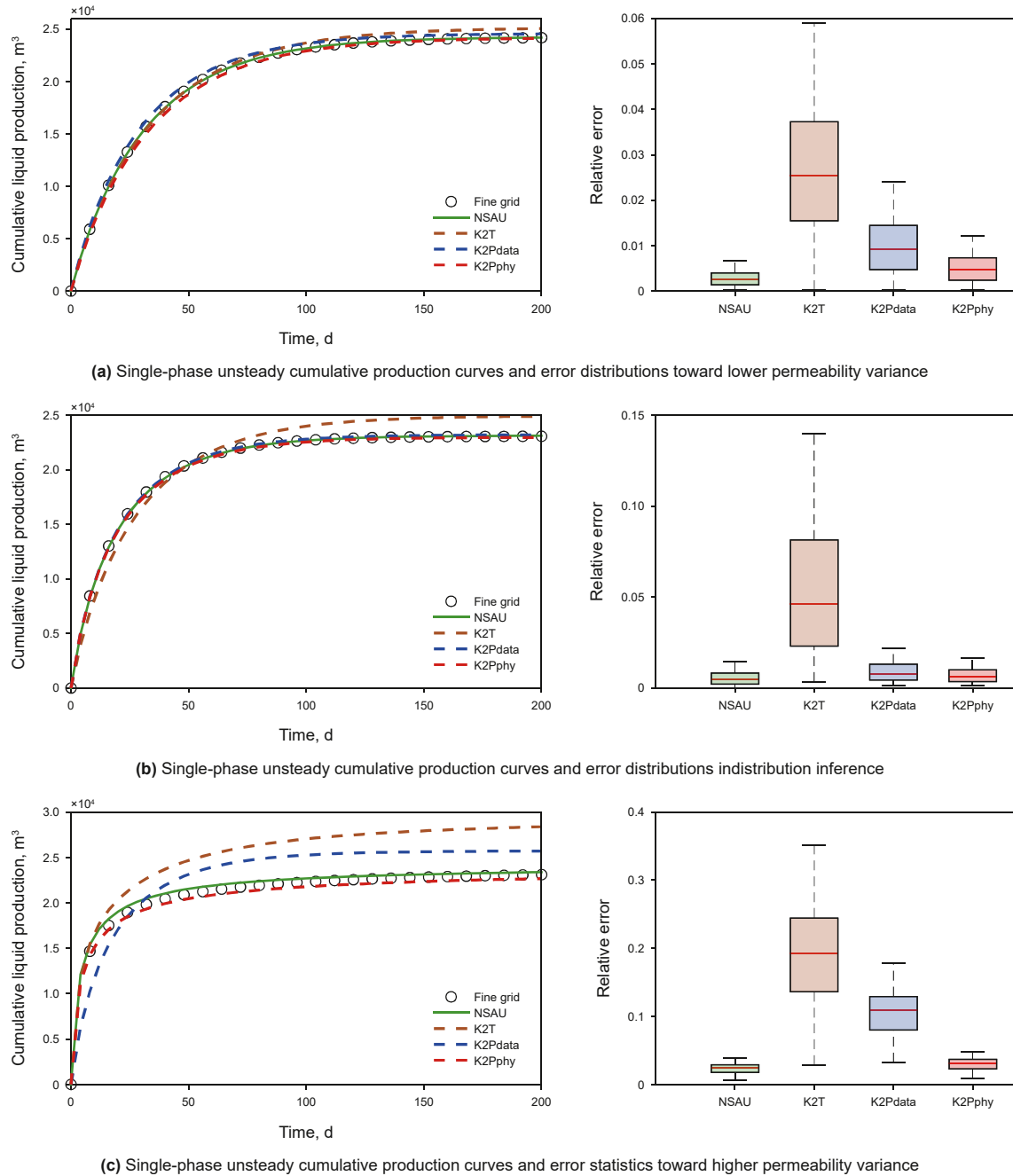


Fig. 16. Single-phase unsteady cumulative production curves and error distributions under different permeability variances in the synthetic CPG model.

largest oscillation amplitude, which suggests a more complex loss landscape, leading to greater convergence challenge. Ultimately, K2Pphy converges to the lowest loss value, implying the extraction of more complex patterns.

After training, predictions are performed on the test set. In Fig. 10, the spatial distribution of common logarithmic transmissibility is visualized, where each line connects the centers of two adjacent grids, and higher transmissibility corresponds to thicker and redder lines. In Fig. 13, the spatial distribution of single-phase steady-state pressure is presented, which also corresponds to the potential/energy fields used to guide upscaling in NSAU, K2Pdata, and K2Pphy. From Figs. 10 and 11, it can be observed that the transmissibility and pressure distributions of K2Pdata and K2Pphy are more consistent with those of NSAU and show better

agreement with the fine-grid transmissibility. In contrast, the K2T result exhibits obvious spatial distribution inconsistencies. We compute the mean value of D_{error} between the NN-assisted methods and NSAU across the entire test set, with K2Pphy achieving the smallest numerical error (mean relative errors with respect to NSAU on the test set; K2T: 0.0311, K2Pdata: 0.0088, K2Pphy: 0.0034). In Fig. 11, it can be observed that the K2T method shows a clear and smooth aggregation of prediction errors in the high-permeability channel between wells P6 and P13, indicating a possible systematic deficiency in recognizing such features. By contrast, the prediction error distributions of K2Pdata and K2Pphy are more random, without obvious error clustering regions.

Fig. 12 presents pressure results of single-phase unsteady flow simulations. It can be seen that NSAU, K2Pdata, and K2Pphy show

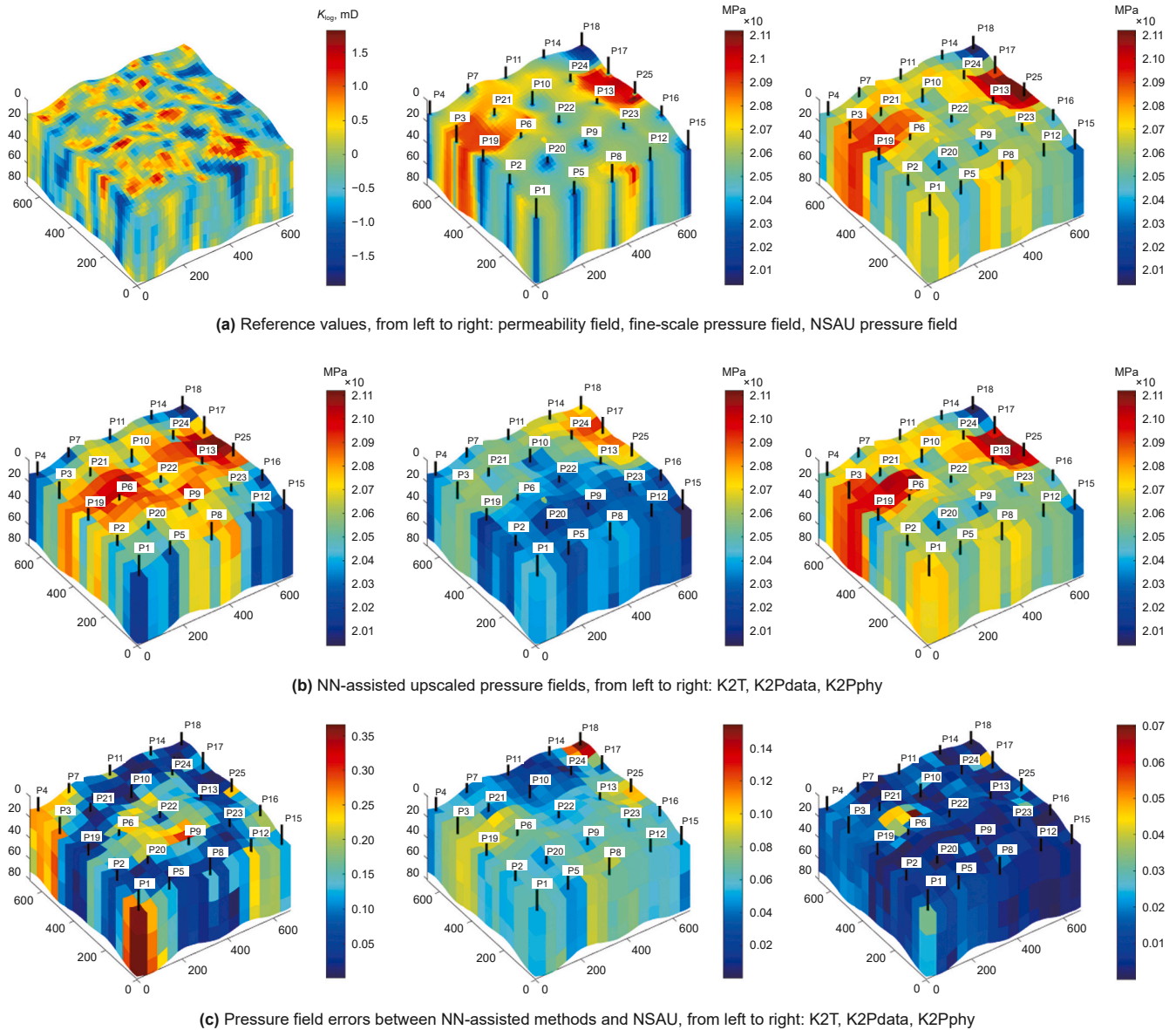


Fig. 17. Single-phase unsteady pressure field distributions toward smaller feature scale under the synthetic CPG model.

good consistency with the fine-scale results and are superior to K2T (K2T: 0.0387, K2Pdata: 0.0191, K2Pphy: 0.0072). However, NSAU, K2Pdata, and K2Pphy still exhibit observable differences in the low-permeability region between wells P5 and P8. The NSAU method tends to characterize this region as significantly high pressure (usually corresponding to relatively lower fluid mobility), which may be the effect of strict constraints in numerical solving. K2Pphy generalizes better than K2Pdata for such features, reflected in higher pressure value, but further improvement may still require enriched sampling of such sparse features. As shown in Fig. 13, the cumulative production curves of NSAU, K2Pdata, and K2Pphy all demonstrate good consistency with the fine-scale results, whereas K2T shows a clear deviation. In the box plots, the prediction errors of NSAU, K2Pdata, and K2Pphy across the entire test set are below 0.03, while the accuracy of K2T is clearly insufficient, with a maximum error approaching 0.15. K2Pphy achieves the accuracy closest to NSAU.

Overall, in model training, K2T converges to a lower loss value than K2Pdata, but its inference generalization on the test set is significantly inferior to that of the K2P methods. This may be

attributed to a fully end-to-end mapping corresponds to a more complex mapping relationship, which also involves more NN parameters (Table 3). Although K2T converges well on the training set, the amount of information and the diversity of features included in the training set may be insufficient to support learning of this more complex mapping. In addition, the significant errors in Fig. 11 indirectly reflect that the K2T method also fails to adequately capture the intermediate mapping relationship from K to p . Among the NN-assisted methods, K2Pphy performs best, particularly in quantitative analyses. This confirms that the PDE-based constraints effectively enhance the extrapolative generalization of NNs to unknown spaces. The predicted pressure fields with better physical consistency also lead to more physical reasonable flow-based upscaling results.

3.2. Extrapolative inference across permeability-variance heterogeneity

The NNs trained in Section 3.1 are set to evaluation mode and are used to infer on the lower permeability-variance and higher

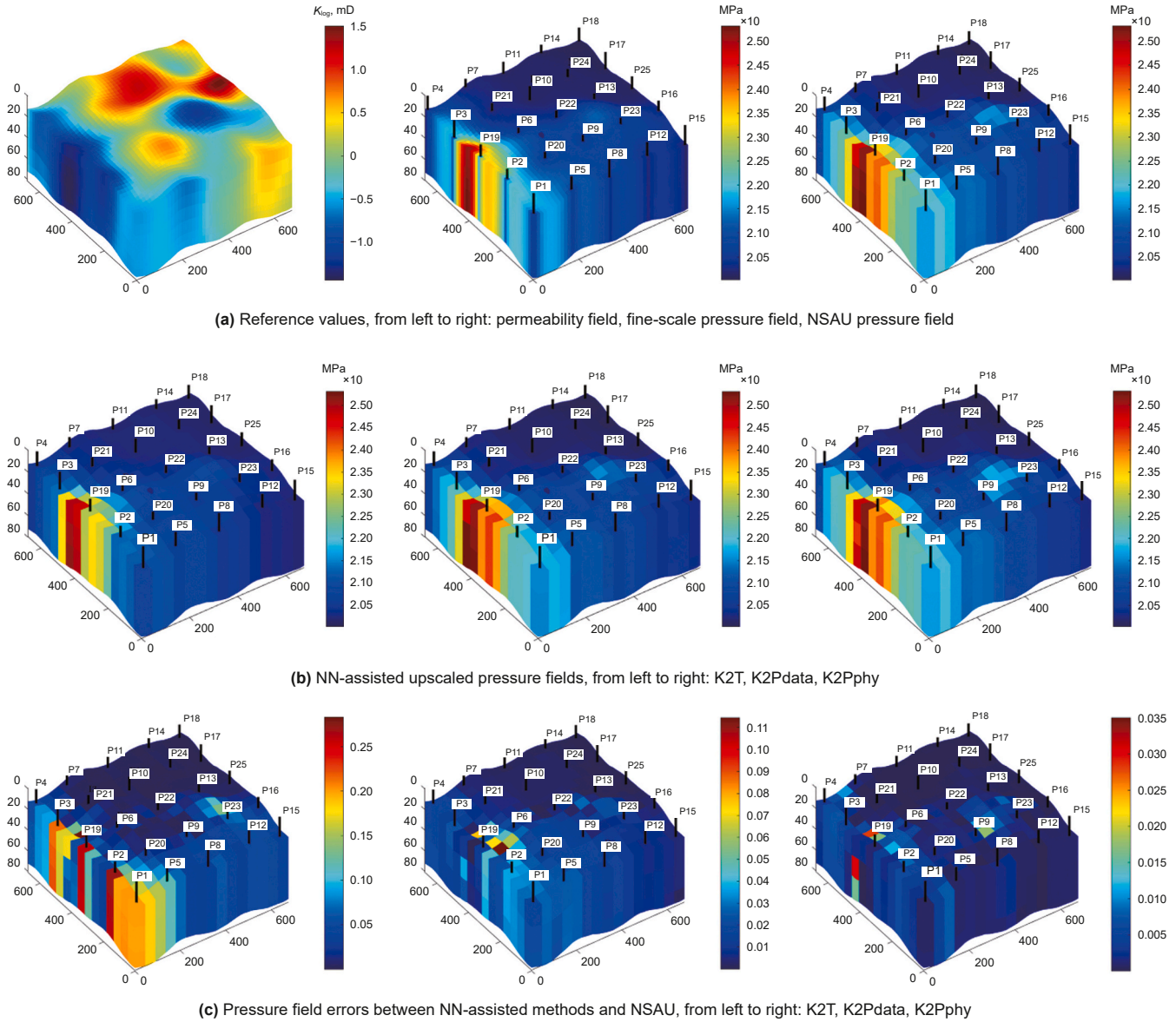


Fig. 18. Single-phase unsteady pressure distributions toward larger scale under the synthetic CPG model.

Table 5

Upscaling errors of inference under different feature scales in the test set of the synthetic model.

Metrics	Cases	NSAU	K2T	K2Pdata	K2Pphy
Average pressure field errors (D_{error})	Toward smaller feature	N/A	0.1413	0.0613	0.0294
	In-distribution inference	N/A	0.0387	0.0191	0.0072
	Toward larger feature	N/A	0.1096	0.0461	0.0142
Mean value of final step relative error of $Q_{1\text{phase}}$	Toward smaller feature	0.0144	0.0776	0.0317	0.0241
	In-distribution inference	0.0059	0.0589	0.0113	0.0086
	Toward larger feature	0.0036	0.0312	0.0158	0.0089

permeability-variance cases, respectively. By comparing Figs. 14 and 15 with Fig. 12, it can be observed that the extrapolative predicted pressure field exhibit more spatial inconsistencies with the reference solution, with particularly significant discrepancies in the results of the K2T model. In Fig. 14, when the permeability variance is smaller, the pressure field distribution is more uniform. The K2Pdata and K2Pphy models simulate a high-pressure zone between P6 and P13 corresponding to the low-permeability

region, with K2Pphy showing better consistency with reference. In contrast, the high-pressure zone of K2T is concentrated only around well P6, which differs from the reference solution. In Fig. 15, when the permeability variance is larger, more scattered high-pressure regions appear, corresponding to low-permeability regions. This correspondence is captured in the results of the K2Pdata and K2Pphy but not in the K2T. When the permeability variance is larger, the NSAU method produces a more prominent

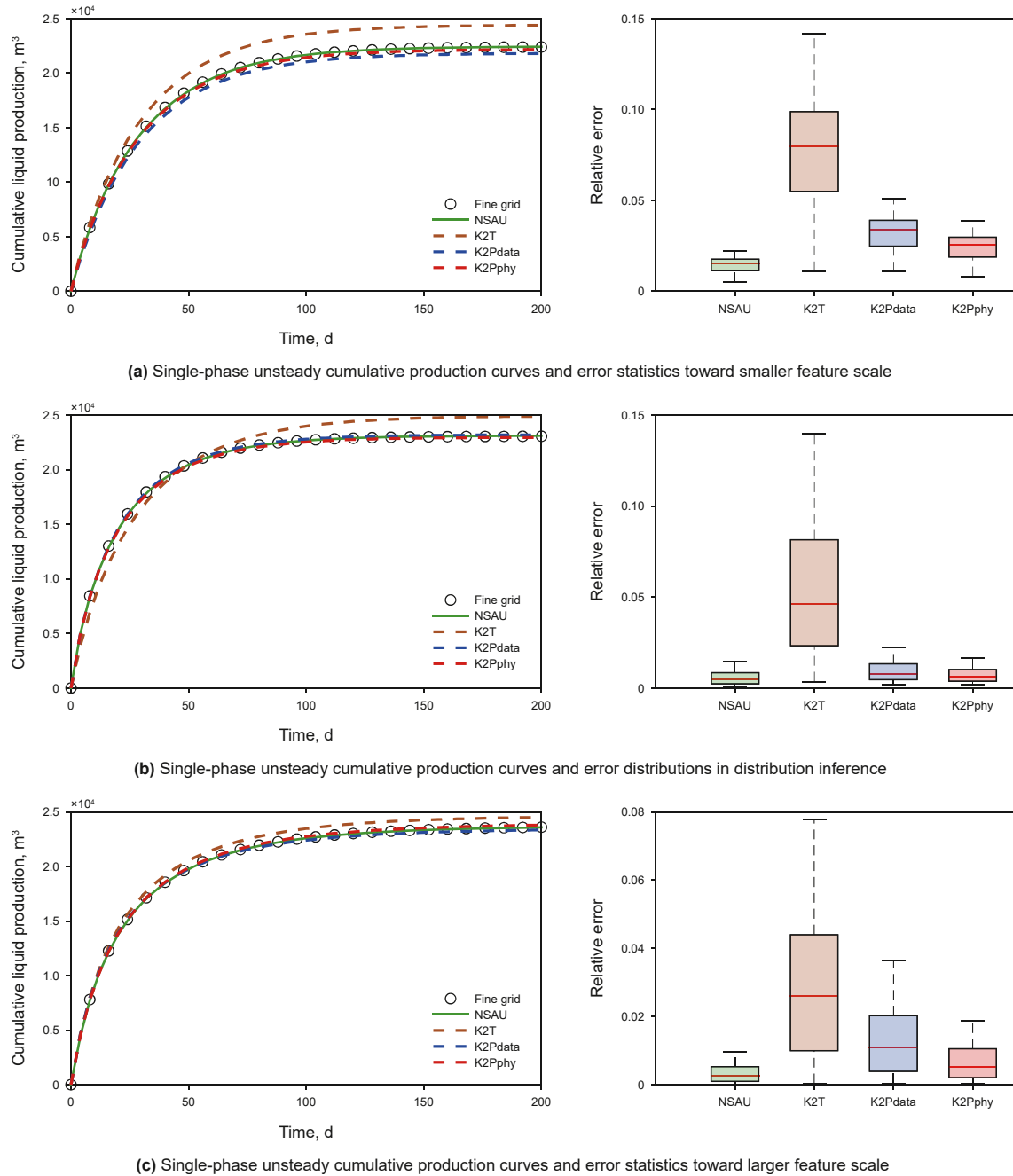


Fig. 19. Single-phase production curves and error distributions under different scales in the synthetic model.

high-pressure zone between P5 and P8, which the NN-assisted methods still fail to reproduce adequately. However, as shown in Fig. 15(c), the error region of K2Pphy is relatively small, mainly near P8, whereas the error regions of K2T and K2Pdata are larger.

In Table 4, it can be observed that the inference errors of the NN-assisted methods on the test set increase when extrapolative inference is performed. In particular, when inferring toward larger permeability variances, the average pressure field error of the K2T method reaches over 0.25, whereas the error of the K2Pphy method remains within 0.05. By comparing Fig. 16(a) and (b), it can be seen that when the permeability variance decreases, the cumulative production curve of the K2T method shows improved consistency with the reference solution. This is because a smaller variance usually indicates that the numerical model is more

homogeneous. Although different upscaling methods vary in their ability to characterize fine-scale permeability distributions, the heterogeneity differences among the upscaled models are reduced, which in turn decreases the response differences of the elastic-drive simulations. From the boxplot in Fig. 16(a), it can be observed that K2T still exhibits relative larger errors compared with K2Pdata and K2Pphy, indicating that the performance improvement resulting from reduced heterogeneity is systematic and arises from the simplification of the simulation problem. When heterogeneity increases, as shown in Fig. 16(c), it can be seen that both K2T and K2Pdata exhibit significant errors, with K2T showing the largest deviation, whereas K2Pphy still maintains relatively better consistency. Similarly, from the boxplot in Fig. 16(c), it can be observed that this performance degradation

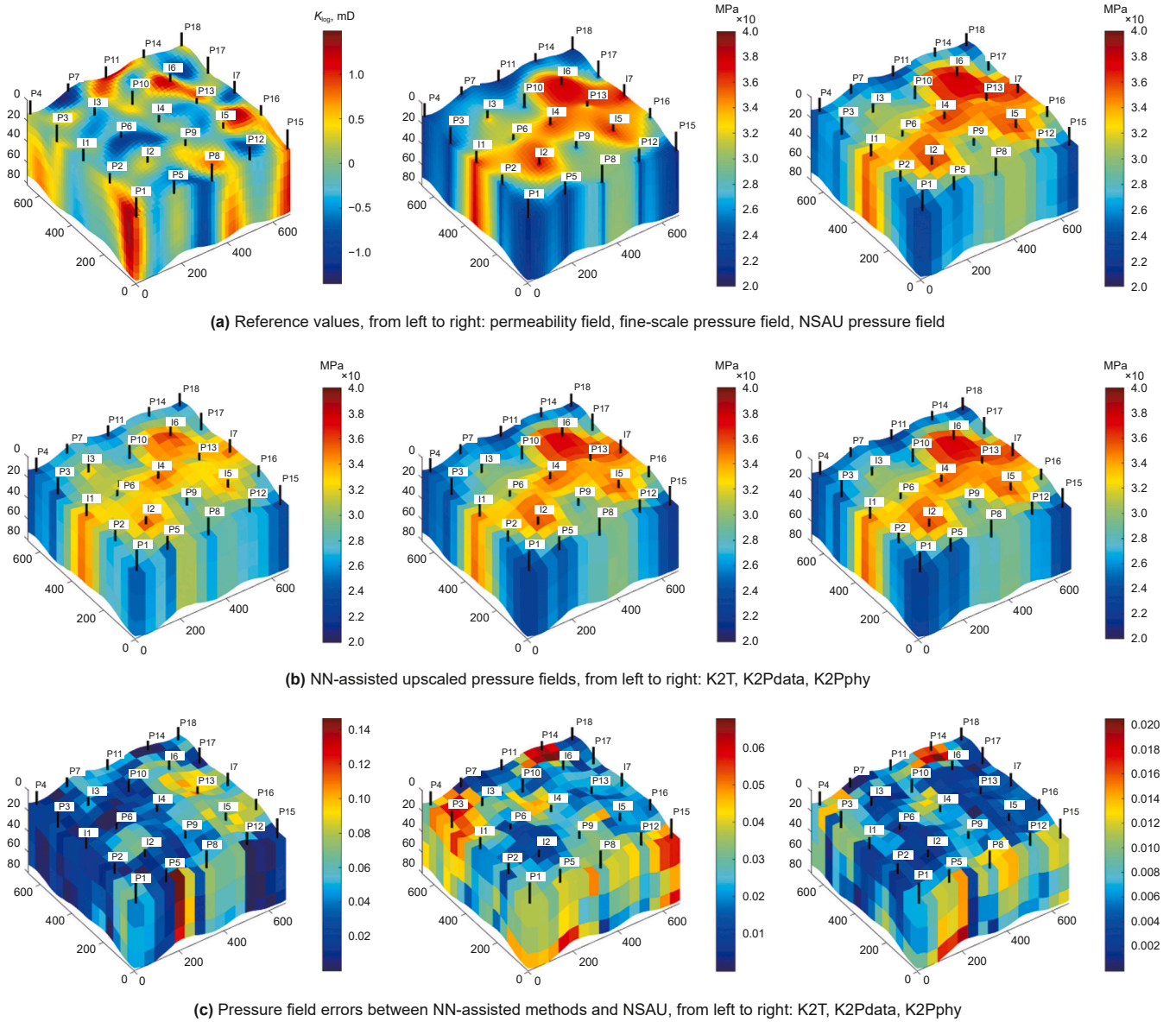


Fig. 20. Two-phase flow pressure field distributions under the synthetic CPG model.

is systematic, while the K2Pphy method still achieves accuracy closest to the NSAU method, with errors below 0.05. Table 4 that K2Pphy consistently achieves the mean value of the final step relative error of $Q_{1\text{phase}}$ closest to that of NSAU. Overall, the K2P mapping strategy enables the NN-assisted upscaling to be more generalized than K2T. The physics-driven training encourages K2Pphy to extract more universal patterns, enhancing its adaptability to the heterogeneity induced by permeability variance varying, thereby achieving best NN-assisted upscaling performance.

3.3. Extrapolative inference across feature-scale heterogeneity

By comparing with Fig. 12, it can be observed that when reasoning toward smaller features (Fig. 17), the simulation results of the K2T and K2Pdata models exhibit significant spatial distribution differences from the reference solutions. The maximum error of K2T exceeds 0.35, while that of K2Pdata exceeds 0.14. A

reasonable explanation is that the reduction of the feature scale introduces more complex local characteristics in the permeability field, which makes the variation patterns of the pressure field more complex and imposes higher requirements on the generalization of the upscaled models. K2T and K2Pdata fail to adequately adapt to the heterogeneity induced by the reduction of the feature-space scale, resulting in poor extrapolative performance. In contrast, although the error of K2Pphy increases, this method still maintains relatively smaller errors and shows spatial distributions more consistent with the reference (Fig. 17(b)). As shown in Fig. 18, when reasoning toward larger features, the high-pressure region is mainly concentrated in the area between P2 and P3, and this characteristic is presented by all the upscaled models. This can be explained that the increase of the feature scale makes more continuous permeability distribution at a large scale, or in other words, the local differences are less pronounced. Consequently, when approximating the response of fine-scale simulations, the demand for local equivalence accuracy is reduced, which

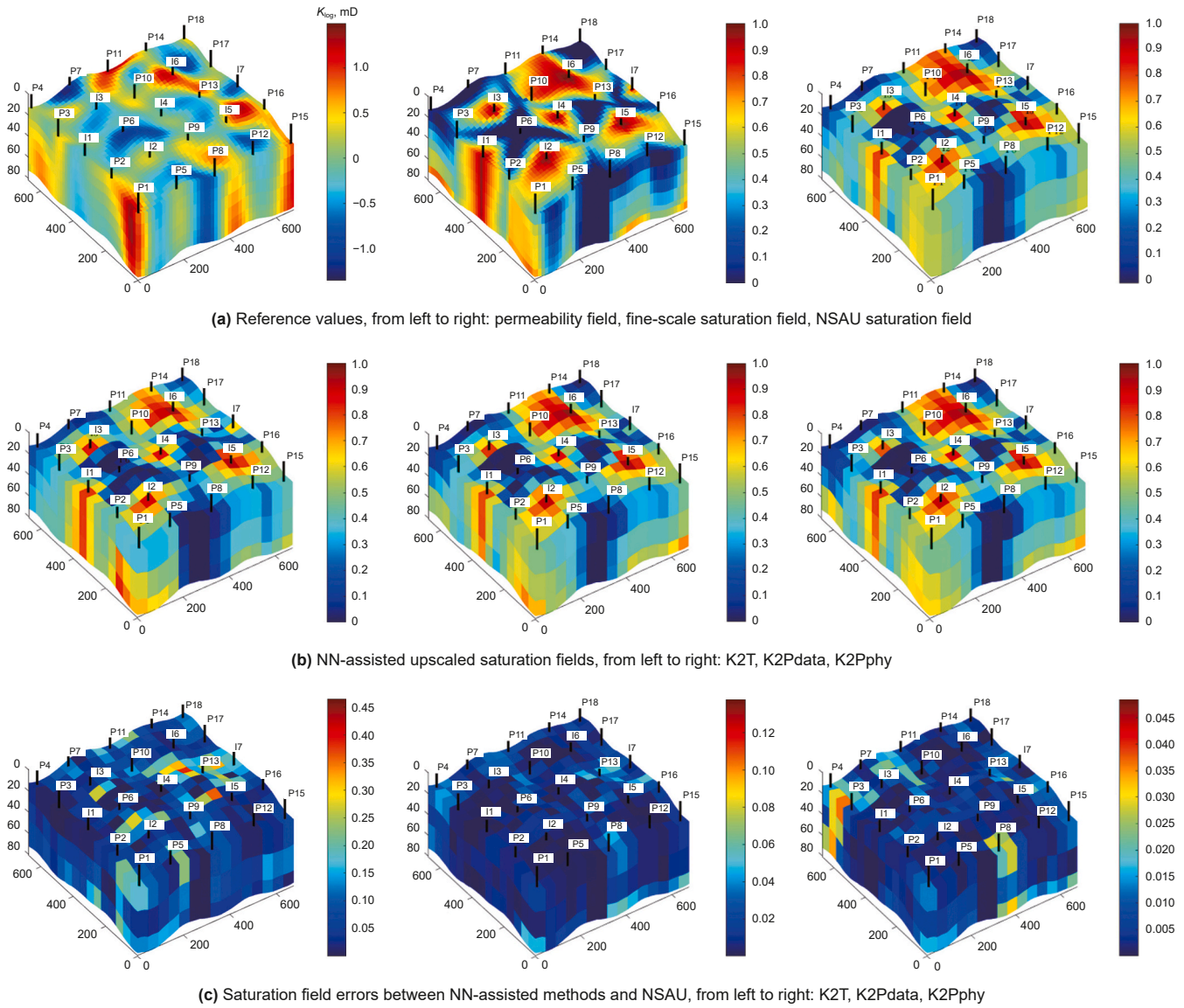


Fig. 21. Two-phase flow saturation field distributions under the synthetic CPG model.

represents a systematic reduction in learning difficulty. Nevertheless, as shown in Fig. 18(c), the error of the K2P method remains smaller, with the K2Pphy method achieving accuracy closest to the NSAU.

In Table 5, the average pressure field errors of different upscaling methods all increase during extrapolation, among which the error of the K2T method expands to above 0.1, while the average error of the K2Pphy method remains within 0.03. By comparing Fig. 19(a) and (b), it can be seen that the reduction of the feature scale leads to a greater deviation between K2T and the reference solution, and the boxplot indicates that this is a systematic increase. Nevertheless, the cumulative production error of the K2Pphy method still remains within 0.05. When the feature scale increases, as shown in Fig. 19(c), the simulation accuracy of different upscaling methods improves, and it is observed that the cumulative production curves become more consistent with the reference solution. This reflects the reduction in upscaling difficulty brought by the larger feature scale, where the influence of local equivalence accuracy differences is diminished. In other words, the upscaled models reproduce similar pressure responses

by capturing the relative permeability differences at the large scale. Nevertheless, the boxplot quantitatively shows that the K2P pathway achieves significantly higher accuracy than K2T, with the average error of the K2Pphy method reduced to within 0.02. It can also be quantitatively observed from Table 5 that K2Pphy consistently achieves a mean value of the final step relative error of $Q_{1\text{phase}}$ closest to that of NSAU, demonstrating the best upscaling accuracy. Overall, among the NN-assisted methods, K2Pphy provides the best upscaling accuracy, whereas K2T performs poorly, especially in cases with greater local heterogeneity. This indicates that optimization of the upscaling pathway and incorporation of physical constraints effectively enhance the extrapolation capability of NN-assisted methods with respect to feature-scale heterogeneity.

3.4. Two-phase flow evaluation

By comparing Fig. 20(a) and (b), it can be seen that the spatial distributions of the pressure fields from K2Pdata and K2Pphy are close to the reference solution, whereas K2T

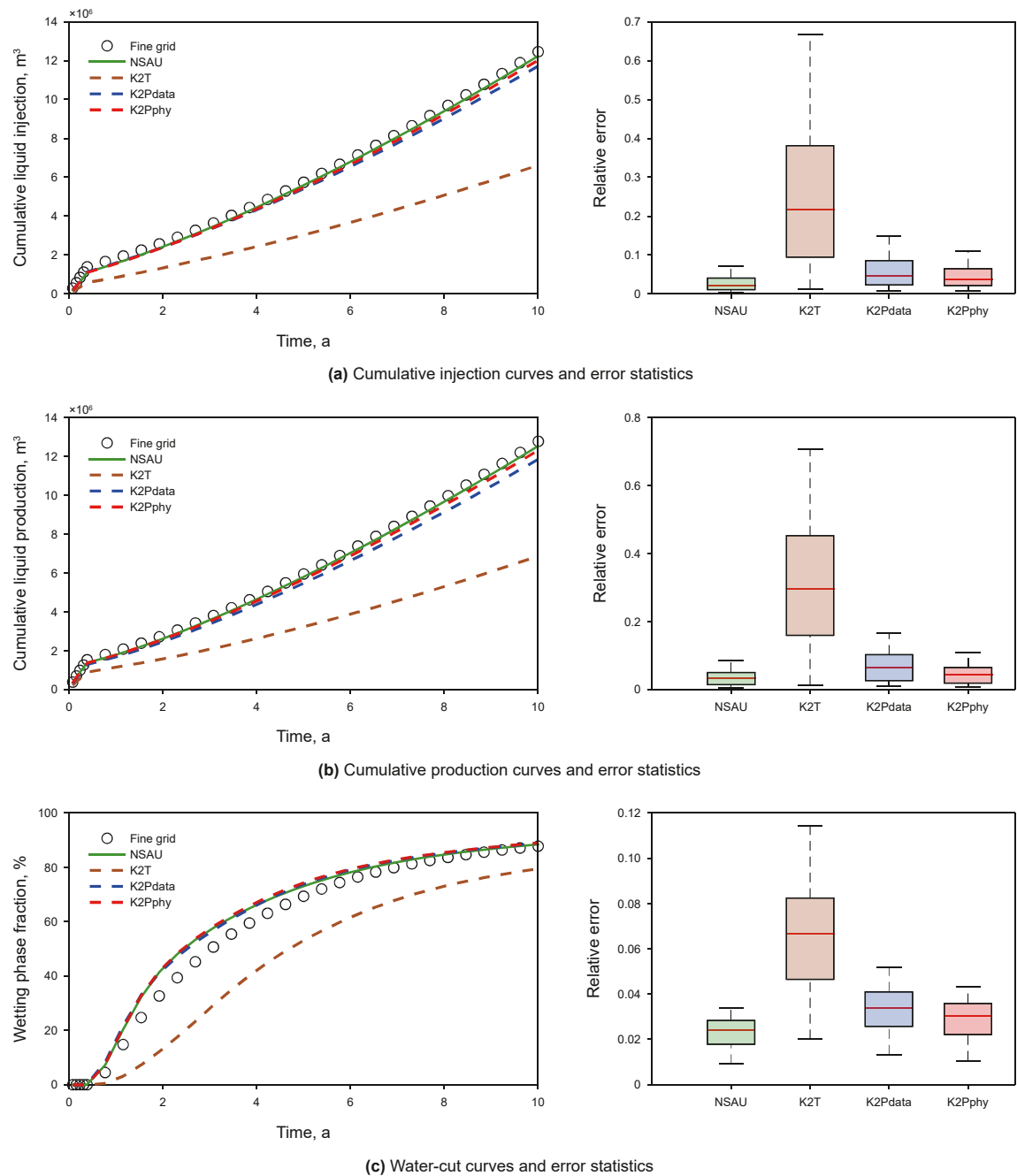


Fig. 22. Two-phase unsteady cumulative injection, cumulative production, and water-cut curves and error distributions under the synthetic CPG model.

Table 6
Upscaling errors of two-phase flow simulation in the test set of the synthetic model.

Metrics	NSAU	K2T	K2Pdata	K2Pphy
Mean value of final step relative error of Q_{inj}	0.0286	0.2648	0.0736	0.0527
Mean value of final step relative error of Q_{pro}	0.0336	0.3098	0.0671	0.0437
Mean value of final step relative error of f_w	0.0228	0.0658	0.0329	0.0288

exhibits obvious differences. From Fig. 20(c), it can be seen that K2Pphy has a lower maximum error. We calculated the average relative errors of the three methods on the test set compared with NSAU (K2T: 0.0587, K2Pdata: 0.0271, K2Pphy: 0.0112). The average error of K2Pphy is around 1%, while that of K2T exceeds 5%. By comparing Fig. 21(a) and (b), it can be seen that K2T,

K2Pdata, and K2Pphy all reproduce the spatial distribution of the saturation field at the final simulation time step in correspondence with the reference, but some differences between K2T and the reference solution can still be observed. In Fig. 21(c), the simulation errors of different NN-assisted methods are clearer, the maximum error of K2T exceeds 0.45,

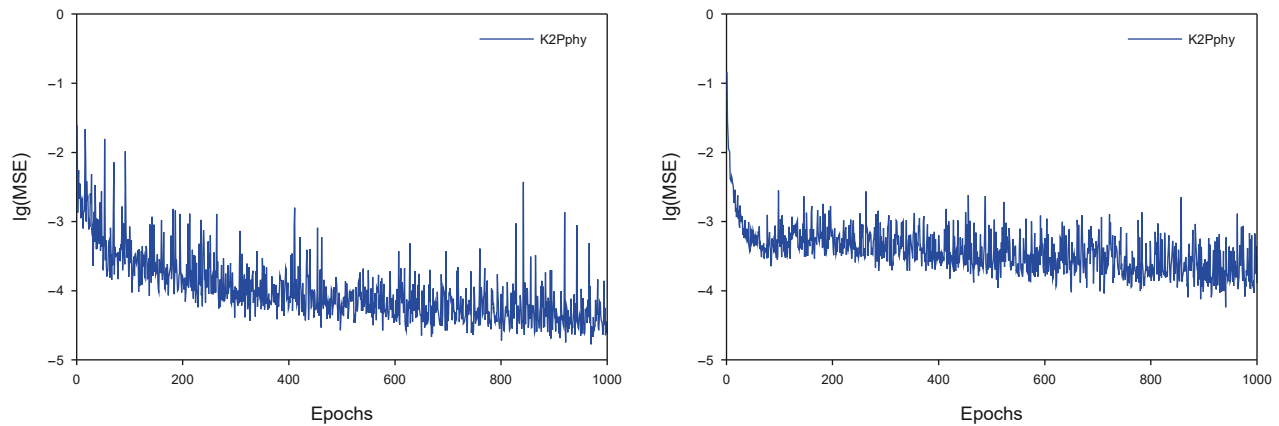


Fig. 23. Training loss curve of K2Pphy in the SAIGUP CPG model (left: train; right: test).

whereas K2Pphy still has the lowest error (<5%). The average relative errors of the three methods on the test set are K2T: 0.1807, K2Pdata: 0.0552, and K2Pphy: 0.0194. Both K2T and K2Pdata exceed 5%, while K2Pphy still performs best.

In Fig. 22, it is evident that the K2T curves show significant deviations from the reference. In particular, the boxplot indicates that its test set error for cumulative injection and cumulative production reach above 0.6 and 0.7, respectively, and the error for water cut exceeds 0.1. The flow rate errors of K2Pdata, K2Pphy, and NSAU also increase; however, their curves remain more consistent with the fine-scale simulation. A reasonable explanation is that the current models only perform single-phase upscaling and insufficiently characterize the relative mobility of two phases, leading to deviations in the simulated fluxes. NSAU, K2Pdata, and K2Pphy provide more accurate

characterization of the overall flow capacity, which results in less degradation of accuracy. In the simulation of water cut, the simulated curves of NSAU, K2Pdata, and K2Pphy also show noticeable shifts relative to the high-resolution solution due to insufficient characterization of the phase front, particularly during the first five years. Nevertheless, the final water cut of NSAU, K2Pdata, and K2Pphy remains closer to the exact solution, with the K2Pphy method achieving accuracy most comparable to NSAU and maintaining a deviation within 0.05 relative to the fine-scale solution. It can also be observed from Table 6 that K2Pphy achieves higher accuracy than K2T and K2Pdata, and is closest to NSAU. Overall, the two-phase flow simulations further demonstrate the advantage of the K2P pathway over K2T and highlight the mitigation of error amplification brought by the embedding of physical constraints in K2Pphy.

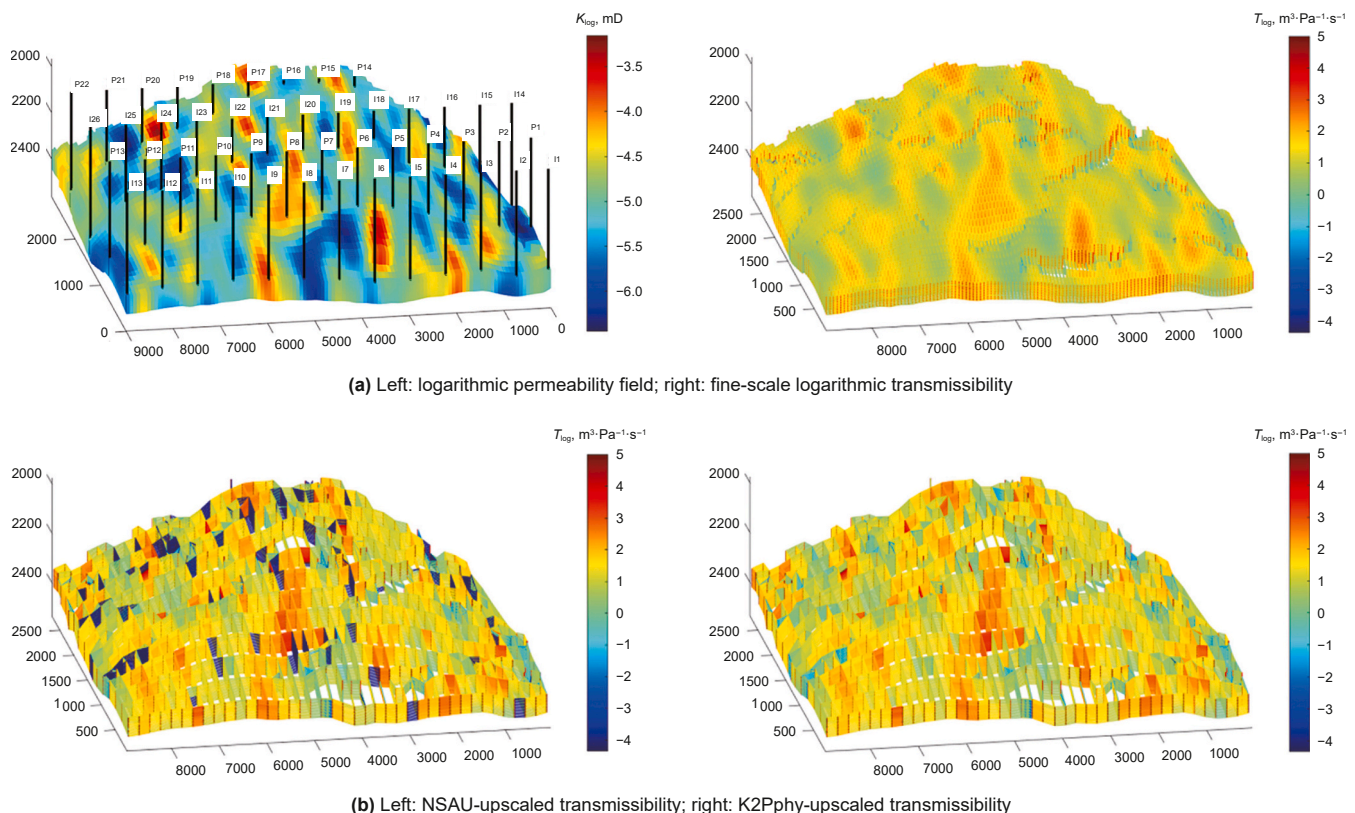


Fig. 24. Comparison of transmissibility spatial distributions of a test sample under the SAIGUP CPG model.

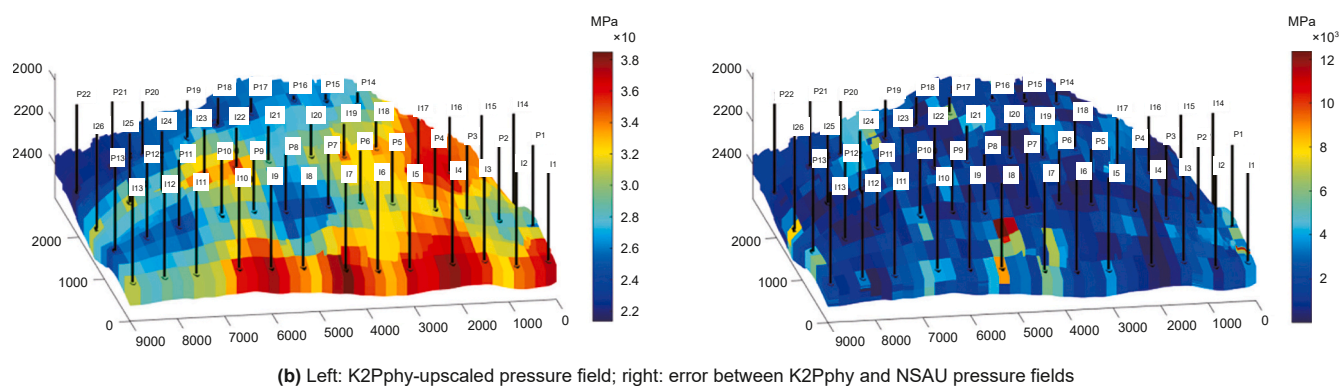
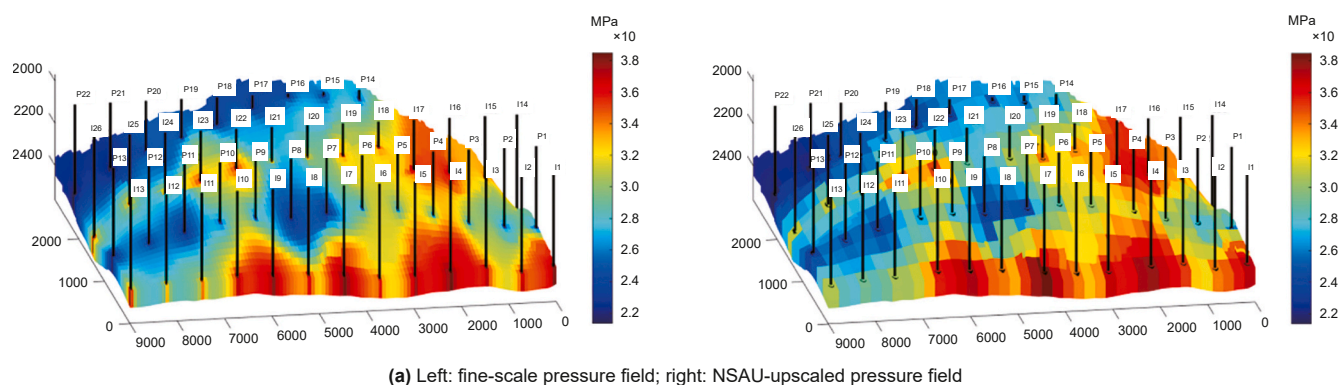


Fig. 25. Comparison of single-phase steady-state pressure of a test sample under the SAIGUP CPG model.

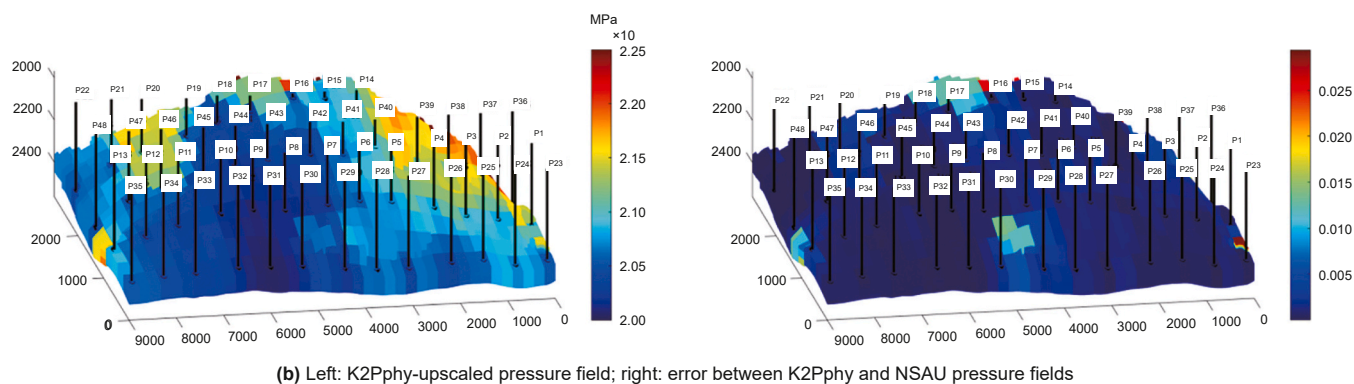
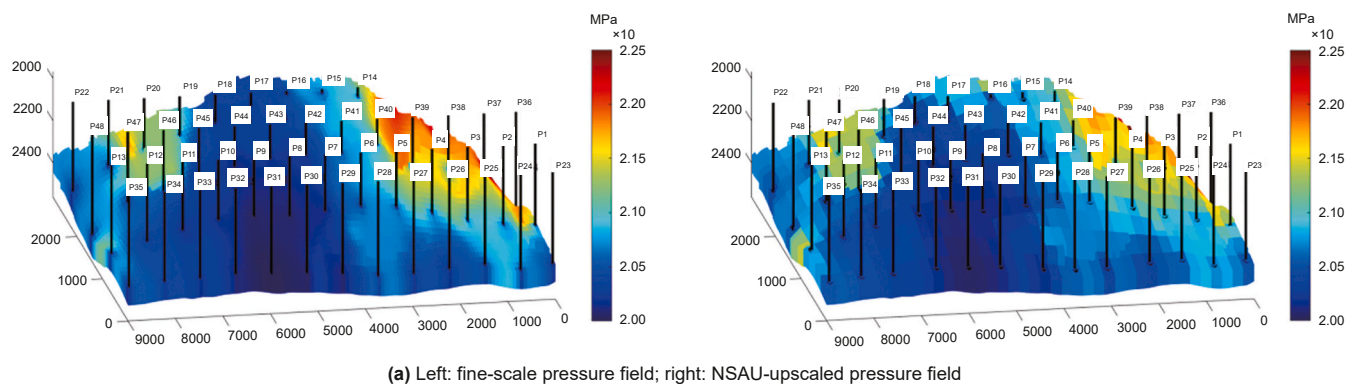


Fig. 26. Single-phase unsteady pressure of a test sample under the SAIGUP CPG model.

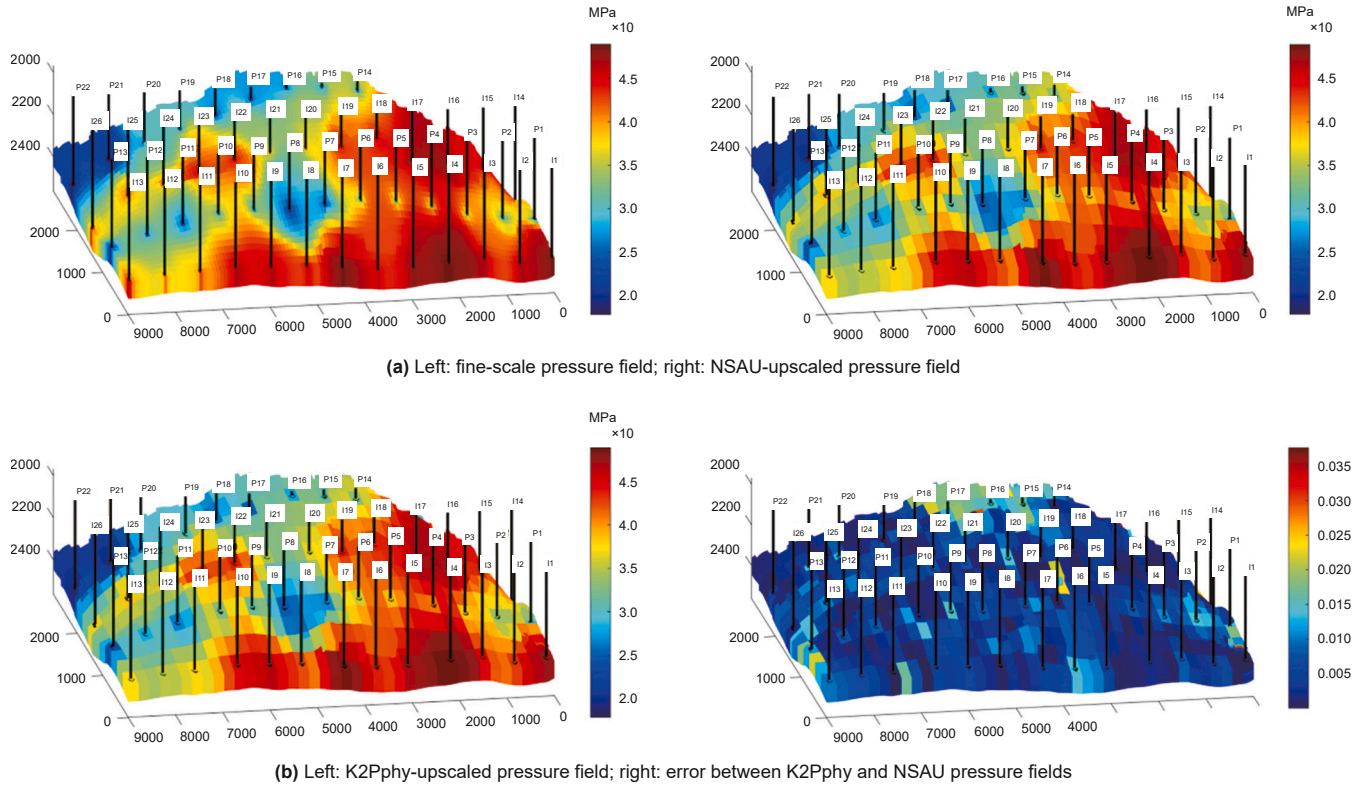


Fig. 27. Two-phase pressure distributions of a test sample under the SAIGUP CPG model.

3.5. Application on the SAIGUP model

After Section 3.1 to Section 3.4, the K2Pphy method, which performs best, is selected to upscale the authentic SAIGUP model,

and it is compared with the fine-scale model and the NSAU method to verify its applicability to more complex CPG models. As shown in Fig. 23, the loss curve exhibits relatively strong oscillations during training, but the overall trend continues to decrease and eventually

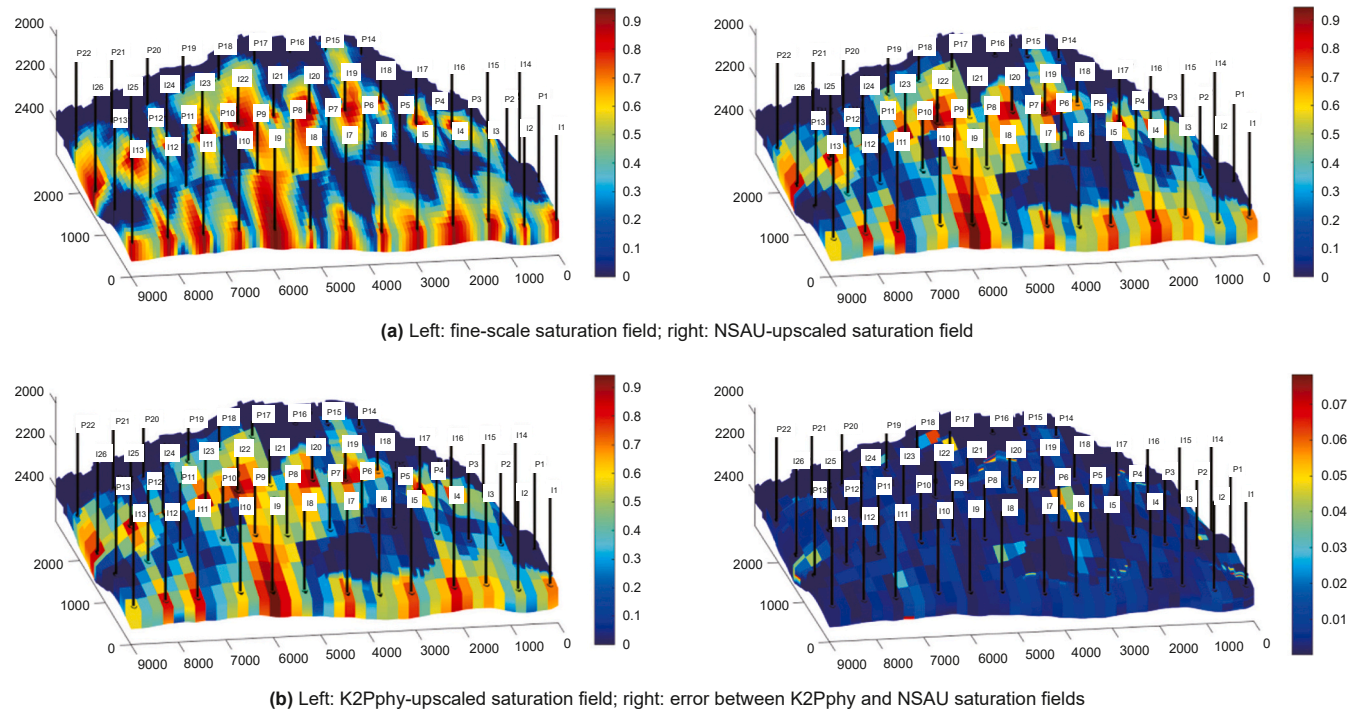


Fig. 28. Two-phase saturation distributions of a test sample under the SAIGUP CPG model.

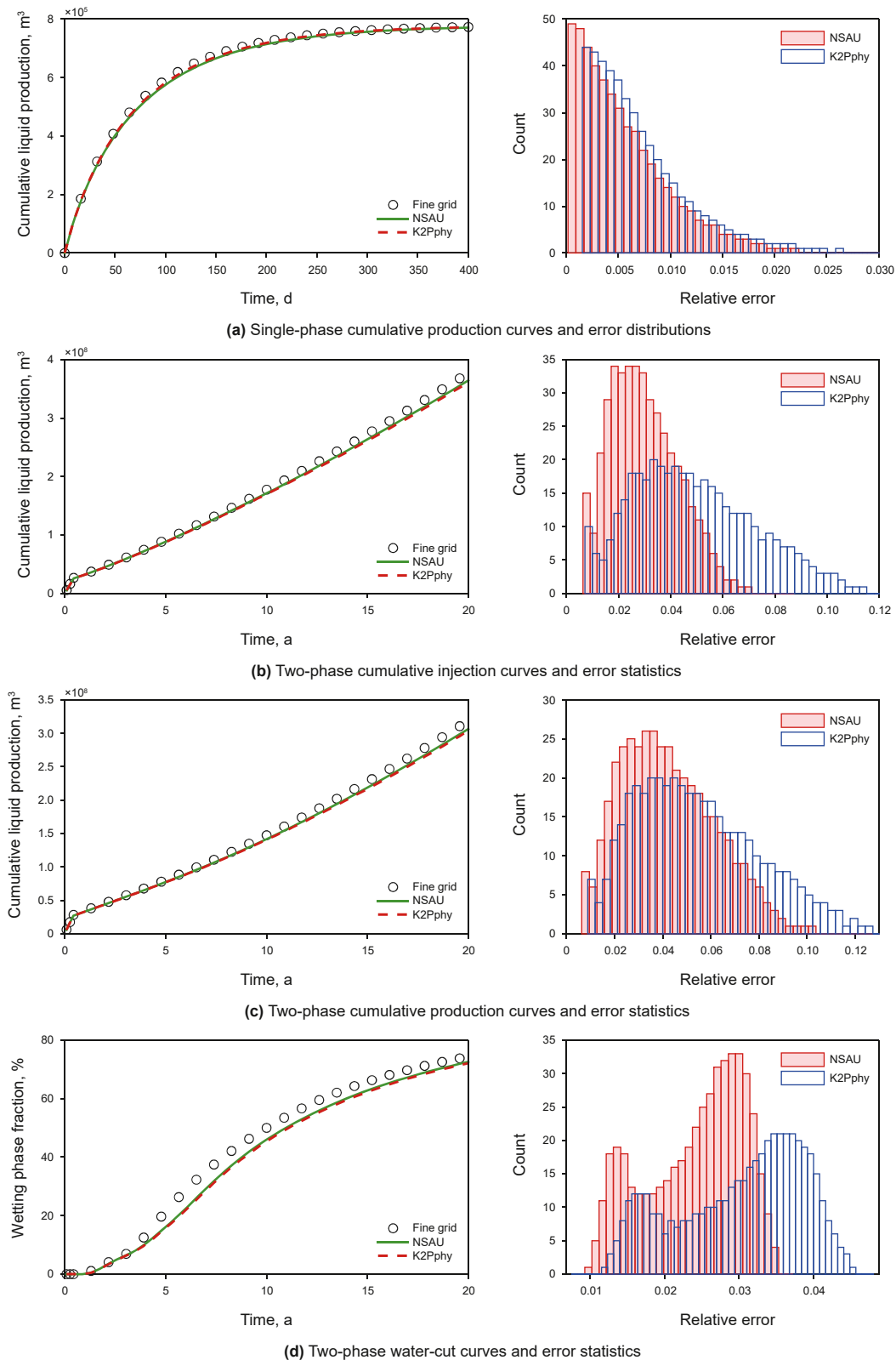


Fig. 29. Comparison of production curves and error distributions of the SAIGUP model.

drops below 1.0×10^{-4} . After training, Fig. 24 shows that the upscaling results of both NSAU and K2Pphy correspond well to the fine-scale model, particularly in regions with higher transmissibility (red regions), which match the high-permeability zones

of the fine-scale model. It is noteworthy that in the upscaling results of NSAU, there are some regions with extremely low conductivity (dark blue), which are not captured in the K2Pphy results. The transmissibility distribution obtained by K2Pphy is overall

Table 7
Upscaling errors in the test set of the SAIGUP CPG model.

Metrics	NSAU	K2Pphy
Mean value of final step relative error of $Q_{1\text{phase}}$	0.006	0.0086
Mean value of final step relative error of Q_{inj}	0.0303	0.051
Mean value of final step relative error of Q_{pro}	0.0433	0.0561
Mean value of final step relative error of f_w	0.024	0.0301

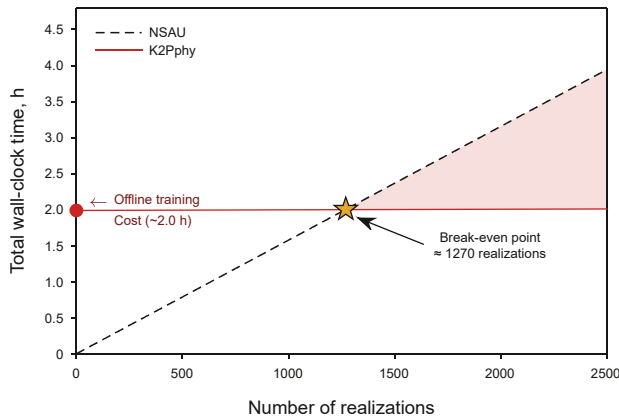


Fig. 30. Computational cost comparison between NSAU and K2Pphy under the SAIGUP CPG model.

smoother. Such discontinuous sharp fronts are computed based on the accurate fine-scale pressure numerical solution under strict flow-equivalence constraints, and they are related to complex inter-block connections and local permeability contrasts.

As shown in Figs. 25–28, the predicted flow fields of K2Pphy are very similar to the spatial distribution of the reference solution, and the average relative error between K2Pphy and NSAU is within 0.03 (single-phase steady-state pressure: 0.0049; single-phase unsteady pressure: 0.0117; two-phase pressure: 0.0151; two-phase saturation: 0.0286). However, the error maps still reveal some differences: in Fig. 25(b), the error distribution between K2Pphy and NSAU is random and uniform, but relatively higher errors appear between the line-drive wells I7, I8 and P7, P8; in Fig. 26(b), two prominent error regions are observed, one of which is similar to the high-error region in the single-phase steady-state error, while another significant error region is located above the wells P15 to P18, which also appears as an error cluster in the two-phase pressure error of Fig. 27(b); in Fig. 28, the spatial distribution of the saturation predicted by K2Pphy and NSAU is generally consistent, but Fig. 28(b) shows a few local regions with larger error values. The upscaling results of K2Pphy do not fully capture all sparse features. This is related to the NN parameter space modeling and nonlinear learning characteristics, which are more suited to approximating functions or functional mappings in a continuous and smooth manner, thereby naturally filtering out some sparse features. If one aims to capture these sparse features, increasing the number of related samples is usually an effective option, although this implies higher computational cost. Therefore, a trade-off between speed and accuracy generally needs to be made depending on specific problems and experimental observations. As shown in Fig. 29(a), the single-phase unsteady cumulative production curves of K2Pphy and NSAU match reference solutions well, and the statistical distribution of final cumulative production errors on the test set is close, less than 0.03. In Fig. 29(b) and (c), K2Pphy achieves production simulation accuracy comparable to NSAU, but both methods show observable decline in accuracy compared with the fine-scale

model. In Fig. 29(d), the water-cut curves of K2Pphy and NSAU both exhibit deviations between the 5th and 15th years, but gradually converge to the reference solution at the final time step, with the test-set final water cut error less than 0.05. These two-phase simulation errors are mainly caused by insufficient characterization of two-phase relative mobility, which needs to be further considered in future works by involving two-phase parameters upscaling. It can be observed from Table 7 that the upscaling error of K2Pphy is higher than that of NSAU, but both consistently remain within the same order of magnitude. Overall, K2Pphy achieves upscaling accuracy comparable to that of NSAU.

To quantitatively assess the computational efficiency, we conducted a wall-clock time comparison between the conventional NSAU and the proposed K2Pphy method. As illustrated in Fig. 30, the total computational cost is plotted against the number of realizations. Although K2Pphy incurs an initial offline training cost of approximately 2.0 h (intercept of the red curve), its marginal cost for online inference is negligible (~0.03 s per realization) compared to the linear cost accumulation of NSAU (~5.7 s per realization). The break-even point occurs at approximately 1270 realizations. Beyond this point, the methodology exhibits a significant advantage. For instance, in a workflow requiring 1900 realizations, K2Pphy saves over 1 h of computational time compared to NSAU. This efficiency gain becomes increasingly pronounced in large-scale ensemble-based tasks, K2Pphy will enable more realizations to be performed under limited computational resources.

4. Conclusions

This study proposes a physics-driven neural network (NN)-assisted global flow-based upscaling method for the rapid upscaling of reservoir models on corner-point grids (CPGs). A finite-volume physics-informed Fourier neural operator is employed to establish a permeability-to-pressure mapping (K2P) for CPG models. The training process does not require labeled data, thereby avoiding expensive numerical simulation costs, and it maintains flux continuity across different grid blocks during inference. By coupling the trained NN with a numerical program for explicit flow-equivalence calculations, the transmissibility and well indices on the coarse grid strictly follow the flow-equivalence relationships with fine-scale, which enables natural applicability to the CPG models. The proposed physics-driven upscaling pathway incorporates physical constraints in both the pressure solution and equivalent calculation stages to enhance inference accuracy. The simulation in the upscaling workflow is transformed into a fast NN forward inference to improve computational efficiency.

In the case study, several NN-assisted methods, including data-driven permeability-to-transmissibility (K2T), data-driven permeability-to-pressure (K2Pdata), and physics-driven permeability-to-pressure (K2Pphy), are compared with the numerical simulation-assisted method (NSAU) and the fine-scale solution. Several key findings confirm the superiority and robustness of K2Pphy. (1) Physical consistency and upscaling accuracy: In the synthetic model, K2Pphy achieves pressure prediction accuracy closest to NSAU, as well as the closest coarse-scale simulation accuracy. This indicates that incorporating physical consistency in the two stages effectively enhances the generalization capability to unknown parameter spaces. (2) Extrapolation generalization capability: When extrapolated to heterogeneous parameter spaces with different permeability variances ($0.1\times$ and $10\times$) and different feature scales (smaller and larger), the trained K2Pphy achieves significantly better accuracy than K2T and K2Pdata, particularly in scenarios with more complex model responses caused by larger permeability variance and smaller feature scale. This

demonstrates that physical constraints guide the NN to extract more universal patterns, resulting in stronger extrapolation capability. (3) Two-phase simulation adaptability: K2Pphy achieves accuracy comparable to NSAU in two-phase flow simulations, which demonstrates that it possesses sufficient upscaling accuracy to mitigate error propagation. (4) Complex CPG model adaptability and computational efficiency. In the application to the SAIGUP model, K2Pphy achieves coarse-scale accuracy comparable to NSAU while exhibiting significantly higher computational efficiency. In particular, when more realizations are required, the computational time of numerical simulation continues to increase sharply, whereas K2Pphy maintains comparable accuracy with only about 0.1% of the additional computational cost of the former.

By combining a discretized physics-constrained training approach with global flow-based upscaling theory, this study develops a novel NN-assisted rapid upscaling method applicable to complex CPG models. It is particularly suitable for periodic closed-loop workflows and ensemble-based scenarios that require repeated execution, enabling broader exploration of uncertain variable spaces under limited computational resources. This is the first implementation of NN-assisted global flow-based upscaling for reservoir models based on CPGs. In particular, it is also the first realization of employing NNs as physics-driven surrogates for such complex CPG models. Future work will extend this research. Regarding physical complexity, we aim to incorporate multipoint flux approximation and two-phase flow mechanisms into the framework to enhance its fidelity in handling complex tensor anisotropy and fluid dynamics. In terms of computational efficiency, we plan to investigate the zero-shot super-resolution capability of the neural operator, leveraging training on coarse grids to generalize to fine-scale models, thereby further reducing training costs. To address fully unstructured grids (e.g., tetrahedral grids) where logical tensor structures are absent, we will explore geometry-agnostic architectures, such as DeepONet or Graph Neural Operators, thereby achieving broader applicability beyond corner-point grids.

CRedit authorship contribution statement

Jing-Qi Lin: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xia Yan:** Writing – review & editing, Validation, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Kai Zhang:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Qi Zhang:** Writing – review & editing, Visualization, Supervision, Project administration, Investigation, Formal analysis. **Zhao Zhang:** Writing – review & editing, Validation, Software, Project administration, Investigation, Data curation. **Liming Zhang:** Writing – review & editing, Visualization, Software, Resources, Project administration, Investigation, Data curation. **Pi-Yang Liu:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Formal analysis.

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 supported by the National Key R&D Program of China (No. 2023YFB4104200), the National Natural Science

Foundation of China (Nos. 52474067, 52441411, 52325402, 52004321, and 12131014), National Key R&D Program of China (No. 2023YFB4104200), Natural Science Foundation of Shandong Province, China (No. ZR2024ME005), Fundamental Research Funds for the Central Universities (Nos. 25CX02025A and 21CX06031A), Youth Innovation and Technology Support Program for Higher Education Institutions of Shandong Province, China (No. 2022KJ070).

Appendix A

This appendix introduces the unsteady single-phase flow problem, including the governing equations, numerical model, and parameter settings. This simulation is used to qualify the upscaling results of Cases 1, 2, 3, and 5. The governing equation for slightly compressible single-phase flow is

$$\frac{\partial}{\partial t}(\phi\rho) + \nabla \cdot (\rho\mathbf{v}_f) = \rho q \quad (\text{A.1})$$

where ϕ denotes porosity. The velocity $\mathbf{v}$ of the unsteady single-phase fluid is defined as

$$\mathbf{v}_f = -\frac{1}{\mu}K(\nabla P - \rho g \nabla D) \quad (\text{A.2})$$

where g is gravitational acceleration, and D is the depth relative to the reference surface. The compressibility of fluid and rock is defined as

$$\rho = \rho_0 \exp [C_f(P - P_0)] \quad (\text{A.3})$$

$$\phi = \phi_0 \exp [C_r(P - P_0)] \quad (\text{A.4})$$

where ρ_0 is the reference density, set to 1000 kg/m³; C_f is the fluid compressibility coefficient, set to 4.5×10^{-9} Pa⁻¹; ϕ_0 is the reference porosity, set to 0.1; C_r is the rock compressibility coefficient, set to 1.0×10^{-11} Pa⁻¹; P_0 is the reference pressure, set to 30 MPa. The source/sink term is still evaluated using Eq. (8).

After FVM discretization, the above model can be solved numerically, and all computations are performed using an in-house simulator (Yan et al., 2016). All grid, permeability, and fluid viscosity settings are consistent with those in Section 2.4. The reservoir pressure is set to 30 MPa, and all injection wells are converted to production wells, with the BHP uniformly set to 20 MPa. This driving mechanism, based on the compression and release of energy from the rock skeleton and pore fluid, is commonly referred to as elastic drive. The synthetic model is simulated for 200 days, and the SAIGUP model is simulated for 400 days.

Appendix B

This appendix introduces the governing equations, numerical model, and parameter settings for slightly compressible two-phase flow, which are used to qualify the upscaling results of Case 4 and Case 5. In two-phase flow, strictly speaking, relative permeability upscaling is also required; nevertheless, in this study, no relative permeability partitioning is performed. In addition, geometric effects may cause differences in relative permeability among different regions, but numerical tests show that the accuracy errors of the models used in this study before and after relative permeability upscaling are small and do not affect the main experimental conclusions. The governing equation for slightly compressible two-phase flow is given as

$$\frac{\partial}{\partial t}(\phi \rho_{\beta} S_{\beta}) + \nabla \cdot (\rho_{\beta} \mathbf{v}_{\beta}) = \rho_{\beta} q_{\beta} \quad (\text{B.1})$$

where β denotes the phase, which can be the wetting phase w or the non-wetting phase n (e.g., water phase and oil phase). The velocity $\mathbf{v}_{\beta}$ is defined as

$$\mathbf{v}_{\beta} = -\frac{k_{r\beta}}{\mu_{\beta}} K(\nabla P_{\beta} - \rho_{\beta} g \nabla D) \quad (\text{B.2})$$

where k_r is the relative permeability, and D is the depth relative to the reference surface. Porosity compressibility is considered through Eq. (A.3), and phase compressibility is considered through

$$\rho_{\beta} = \rho_{\beta 0} \exp [C_{\beta}(P - P_0)] \quad (\text{B.3})$$

where $\rho_{\beta 0}$ is the reference density of phase β ; C_{β} is the compressibility coefficient of phase β . Additional equations are defined as

$$S_w + S_n = 1 \quad (\text{B.4})$$

$$P_c = P_n - P_w \quad (\text{B.5})$$

where P_c is the capillary pressure, which is neglected in this study, i.e., $P_n = P_w$. The source/sink term is defined as

$$q_{\beta} = \lambda_{\beta} \frac{2\pi Kh}{\ln(r_e/r_w) + s} (P - P_{wf}) \quad (\text{B.6})$$

where λ_{β} is the mobility of phase β , defined as $\lambda_{\beta} = k_r / \mu_{\beta}$.

The model can be solved numerically, and all computations are performed using (Yan et al., 2016). The grid, well configuration, permeability, and other parameter settings of the two-phase numerical model are consistent with those in Section 2.4. Porosity is set to 0.1, and rock compressibility is set to $1.0 \times 10^{-11} \text{ Pa}^{-1}$. The relative permeability curves are shown in Fig. B.1, and the detailed fluid properties are listed in Table B.1. The reservoir pressure is set to 30 MPa, the BHP of production wells is set to 20 MPa, and the BHP of injection wells is set to 40 MPa. The synthetic model is simulated for 10 year, and the SAIGUP model is simulated for 20 years.

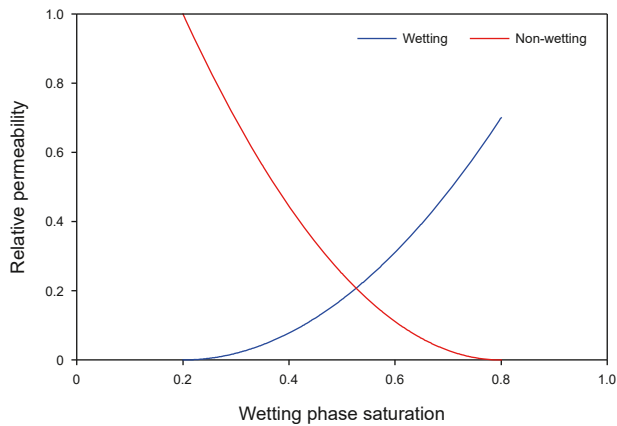


Fig. B.1. Relative permeability curves for the two-phase flow problem.

Table B.1

Two-phase fluid parameter settings.

Parameter	Value	Parameter	Value
Wetting phase density, kg/m ³	1000	Wetting phase compressibility, 1/Pa	1.0×10^{-10}
Non-wetting phase density, kg/m ³	700	Non-wetting phase compressibility, 1/Pa	4.5×10^{-9}
Wetting phase viscosity, mPa·s	1	Irreducible wetting phase saturation	0.2
Non-wetting phase viscosity, mPa·s	5	Residual non-wetting phase saturation	0.2

Appendix C

This appendix presents the sensitivity analysis results for the synthetic model concerning key factors, as summarized in Table C.1 and Fig. C.1. First, regarding permeability variance, the results indicate that while the inference error exhibits a slight upward trend as the variance increases, the sensitivity of all methods (K2T, K2Pdata, and K2Pphy) to this factor remains low. This marginal increase in error is primarily attributed to the inherent escalation in system complexity resulting from stronger heterogeneity, rather than deficiencies in the network configuration. Second, the sensitivity test for network depth reveals that the error continuously decreases as the number of hidden layers increases from 1 to 3; however, a slight rebound in error is observed when the depth reaches 4 layers, suggesting that a 3-layer structure represents the optimal configuration. Finally, the analysis of the number of training samples demonstrates that the error decreases monotonically with increasing sample size. Notably, beyond 500 samples, the performance gains yield diminishing returns while computational resource consumption inevitably rises. Therefore, a dataset of 500 samples is identified as the near-optimal size that balances accuracy with efficiency.

Table C.1

Sensitivity analysis results on key factors.

Sensitivity factors	Value	Mean value of flow field D_{error}		
		K2T	K2Pdata	K2Pphy
Permeability variance	0.1	0.03079	0.00875	0.00338
	0.2	0.03098	0.00880	0.00340
	0.3	0.03112	0.00884	0.00341
	0.4	0.03125	0.00887	0.00342
Number of hidden layers	1	0.03385	0.00950	0.00365
	2	0.03190	0.00903	0.00347
	3	0.03112	0.00884	0.00341
	4	0.03189	0.00898	0.00345
Number of training samples	100	0.03771	0.01057	0.00406
	300	0.03343	0.00942	0.00362
	500	0.03112	0.00884	0.00341
	700	0.03055	0.00871	0.00337

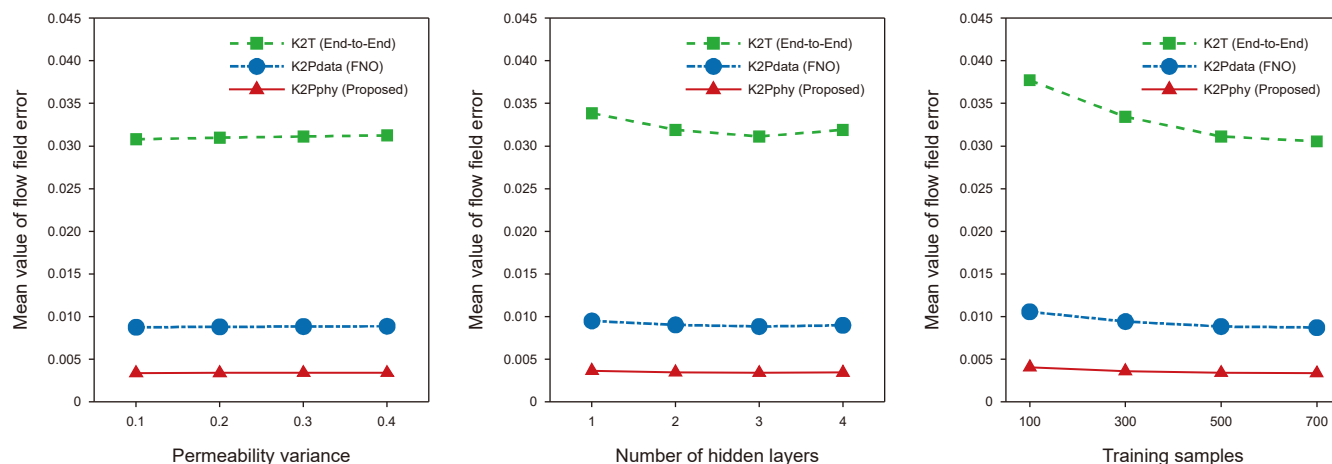


Fig. C.1. Sensitivity analysis curves on key factors (left: sensitivity to permeability variance; middle: sensitivity to network depth; right: sensitivity to number of training samples).

References

- Akiba, T., Sano, S., Yanase, T., et al., 2019. Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631. <https://doi.org/10.1145/3292500.3330701>.
- Alpak, F.O., 2021. Practical implementation of a method for global single-phase flow-based transmissibility upscaling using generic flow boundary conditions and its application on models with non-local heterogeneities. *J. Petrol. Sci. Eng.* 207, 109037. <https://doi.org/10.1016/j.petrol.2021.109037>.
- Banaei, S.M.A., Javid, A.H., Hassani, A.H., 2021. Numerical simulation of groundwater contaminant transport in porous media. *Int. J. Environ. Sci. Technol.* 18 (1), 151–162. <https://doi.org/10.1007/s13762-020-02825-7>.
- Cai, S., Mao, Z., Wang, Z., et al., 2021. Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mech. Sin.* 37 (12), 1727–1738. <https://doi.org/10.1007/s10409-021-01148-1>.
- Cao, Q., Banerjee, R., Gupta, S., et al., 2016. Data-driven production forecasting using machine learning. In: SPE Argentina Exploration and Production of Unconventional Resources Symposium. <https://doi.org/10.2118/180984-MS>.
- Chen, Y., Oliver, D.S., Zhang, D., 2009. Efficient ensemble-based closed-loop production optimization. *SPE J.* 14 (4), 634–645. <https://doi.org/10.2118/112873-PA>.
- Chen, Y., Durlafsky, L.J., 2008. Ensemble-level upscaling for efficient estimation of fine-scale production statistics. *SPE J.* 13 (4), 400–411. <https://doi.org/10.2118/106086-PA>.
- Chen, Z., Huan, G., Ma, Y., 2006. Computational Methods for Multiphase Flows in Porous Media. SIAM, Philadelphia, PA. <https://doi.org/10.1137/1.9780898718942>.
- Chi, L., Jiang, B., Mu, Y., 2020. Fast Fourier convolution. In: NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems, pp. 4479–4488. https://proceedings.neurips.cc/paper_files/paper/2020/hash/2fd5441ec6cfab47e32164d5624269b1-Abstract.html.
- Dagan, G., Fiori, A., Jankovic, I., 2013. Upscaling of flow in heterogeneous porous formations: Critical examination and issues of principle. *Adv. Water Resour.* 51, 67–85. <https://doi.org/10.1016/j.advwatres.2011.12.017>.
- Devlin, J., Chang, M.W., Lee, K., et al., 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4171–4186. <https://doi.org/10.18653/v1/N19-1423>.
- Donnelly, J., Daneshkhah, A., Abolfathi, S., 2024. Physics-informed neural networks as surrogate models of hydrodynamic simulators. *Sci. Total Environ.* 912, 168814. <https://doi.org/10.1016/j.scitotenv.2023.168814>.
- Durlafsky, L.J., 2005. Upscaling and gridding of fine-scale geological models for flow simulation. In: 8th International Forum on Reservoir Simulation, pp. 1–59.
- Eshaghi, M.S., Anitescu, C., Thombre, M., et al., 2025. Variational physics-informed neural operator (VINO) for solving partial differential equations. *Comput. Methods Appl. Mech. Eng.* 437, 117785. <https://doi.org/10.1016/j.cma.2025.117785>.
- Eymard, R., Gallouët, T., Herbin, R., 2000. Finite volume methods. In: Handbook of Numerical Analysis. Elsevier, Amsterdam, pp. 713–1018. [https://doi.org/10.1016/S1570-8659\(00\)07005-8](https://doi.org/10.1016/S1570-8659(00)07005-8).
- Farmer, C.L., 2002. Upscaling: A review. *Int. J. Numer. Methods Fluid.* 40, 63–78. <https://doi.org/10.1002/fld.267>.
- Feldmann, F., Hagemann, B., Ganzer, L., et al., 2016. Numerical simulation of hydrodynamic and gas mixing processes in underground hydrogen storages. *Environ. Earth Sci.* 75 (16), 1165. <https://doi.org/10.1007/s12665-016-5948-z>.
- Floris, F.J.T., Bush, M.D., Cuypers, M., et al., 2001. Methods for quantifying the uncertainty of production forecasts: A comparative study. *Pet. Geosci.* 7 (Suppl. 1), S87–S96. <https://doi.org/10.1144/petgeo.7.S.S87>.
- Fu, D., Fu, Y., Zhang, Y., et al., 2024. Experimental simulation study on influencing factors of liquid production capacity in heterogeneous water drive reservoirs. *Phys. Fluids* 36 (1), 016612. <https://doi.org/10.1063/5.0184992>.
- Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, pp. 249–256. In: <https://proceedings.mlr.press/v9/glorot10a.html>.
- Harandi, A., Danesh, H., Linka, K., et al., 2025. SPiFOL: A spectral-based physics-informed finite operator learning for prediction of mechanical behavior of microstructures. *J. Mech. Phys. Solid.* 203, 106219. <https://doi.org/10.1016/j.jmps.2025.106219>.
- Hendrycks, D., Gimpel, K., 2016. Gaussian Error Linear Units (GELUs). arXiv: 1606.08415. <https://doi.org/10.48550/arXiv.1606.08415>.
- Holden, L., Nielsen, B.F., 2000. Global upscaling of permeability in heterogeneous reservoirs; the output Least squares (OLS) method. *Transport Porous Media* 40 (2), 115–143. <https://doi.org/10.1023/A:1006657515753>.
- Houlding, S.W., 1994. 3D Geoscience Modeling: Computer Techniques for Geological Characterization. Springer Berlin Heidelberg, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-79012-6>.
- Howell, J.A., Skorstad, A., MacDonald, A., et al., 2008. Sedimentological parameterization of shallow-marine reservoirs. *Pet. Geosci.* 14 (1), 17–34. <https://doi.org/10.1144/1354-079307-787>.
- Islam, M.S., Manzocchi, T., Walsh, J.J., et al., 2021. A comparative study of geometrical upscaling methods for inclusion of fault zone structure into upscaled reservoir simulation models. *Nat. Resour. Res.* 30 (3), 2561–2573. <https://doi.org/10.1007/s11053-021-09832-6>.
- Jansen, J.D., Douma, S.D., Brouwer, D.R., et al., 2009. Closed-loop reservoir management. In: SPE Reservoir Simulation Symposium. <https://doi.org/10.2118/119098-MS>.
- Jiang, S., Durlafsky, L.J., 2023. Use of multifidelity training data and transfer learning for efficient construction of subsurface flow surrogate models. *J. Comput. Phys.* 474, 111800. <https://doi.org/10.1016/j.jcp.2022.111800>.
- Jin, Z.L., Liu, Y., Durlafsky, L.J., 2020. Deep-learning-based surrogate model for reservoir simulation with time-varying well controls. *J. Petrol. Sci. Eng.* 192, 107273. <https://doi.org/10.1016/j.petrol.2020.107273>.
- Ketkar, N., Moolayil, J., 2021. Introduction to PyTorch. In: Ketkar, N., Moolayil, J. (Eds.), Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch. Apress, Berkeley, CA, pp. 27–91. https://doi.org/10.1007/978-1-4842-5364-9_2.
- Khiri, S., Feyzmahdavian, H.R., Johansson, M., 2017. Mini-batch gradient descent: Faster convergence under data sparsity. In: 2017 IEEE 56th Annual Conference on Decision and Control (CDC), pp. 2880–2887. <https://doi.org/10.1109/CDC.2017.8264077>.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv: 1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>.
- Kodama, Y., Ouchi, H., Kurihara, M., 2024. 2-dimensional upscaling using deep learning. In: SPWLA 29th Formation Evaluation Symposium of Japan, SPWLA-JFES-2024-Gd.
- Koric, S., Abueidda, D.W., 2023. Data-driven and physics-informed deep learning operators for solution of heat conduction equation with parametric heat source. *Int. J. Heat Mass Tran.* 203, 123809. <https://doi.org/10.1016/j.jheatmasstransfer.2022.123809>.

- Kostakis, F., Mallison, B.T., Durlflosky, L.J., 2020. Multifidelity framework for uncertainty quantification with multiple quantities of interest. *Comput. Geosci.* 24 (2), 761–773. <https://doi.org/10.1007/s10596-019-9825-1>.
- Krishnamurthy, P., Lun, L., Li, H., et al., 2024. Investigating how geological heterogeneity is modeled impact on large scale CO₂ storage performance prediction. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.5035364>.
- Kumar, A., Ozah, R., Noh, M., et al., 2005. Reservoir simulation of CO₂ storage in deep saline aquifers. *SPE J.* 10 (3), 336–348. <https://doi.org/10.2118/89343-PA>.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>.
- Lee, M., 2023. GELU activation function in deep learning: A comprehensive mathematical analysis and performance. *arXiv:2305.12073*. <https://doi.org/10.48550/arXiv.2305.12073>.
- Li, D., Ionescu, C.L., Muftakhidinov, B., et al., 2021. A new simulation layer optimization and permeability upscaling method for preserving critical reservoir heterogeneity. In: *SPE Annual Caspian Technical Conference*. <https://doi.org/10.2118/207074-MS>.
- Li, Z., Kovachki, N., Azizzadenesheli, K., et al., 2020. Fourier Neural Operator for Parametric Partial Differential Equations. *arXiv:2010.08895*. <https://doi.org/10.48550/arXiv.2010.08895>.
- Li, Z., Zheng, H., Kovachki, N., et al., 2024. Physics-informed neural operator for learning partial differential equations. *ACM/IMS Journal of Data Science* 1 (3), 9. <https://doi.org/10.1145/3648506>.
- Lie, K.A., 2019. *An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST)*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781108591416>.
- Lin, J.Q., Yan, X., Wang, E.Z., et al., 2025. A layer-specific constraint-based enriched physics-informed neural network for solving two-phase flow problems in heterogeneous porous media. *Pet. Sci.* 22 (11), 4714–4735. <https://doi.org/10.1016/j.petsci.2025.07.008>.
- Lin, J., Zhang, K., Zhang, L., et al., 2024. Towards efficient and accurate CO₂ sequestration optimization: integrating hierarchical spatio-temporal information into deep-learning-based surrogate models. *Fuel* 356, 129343. <https://doi.org/10.1016/j.fuel.2023.129343>.
- Lu, L., Jin, P., Pang, G., et al., 2021. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nat. Mach. Intell.* 3 (3), 218–229. <https://doi.org/10.1038/s42256-021-00302-5>.
- Lu, L., Meng, X., Cai, S., et al., 2022. A comprehensive and fair comparison of two neural operators (with practical extensions) based on FAIR data. *Comput. Methods Appl. Mech. Eng.* 393, 114778. <https://doi.org/10.1016/j.cma.2022.114778>.
- Ma, Z., Ou, X., Zhang, B., 2024. Development of a convolutional neural network based geomechanical upscaling technique for heterogeneous geological reservoir. *J. Rock Mech. Geotech. Eng.* 16 (6), 2111–2125. <https://doi.org/10.1016/j.jrmge.2024.02.009>.
- Müller, S., Schüller, L., Zech, A., et al., 2022. GSTools v1.3: A toolbox for geostatistical modelling in Python. *Geosci. Model Dev.* 15 (7), 3161–3182. <https://doi.org/10.5194/gmd-15-3161-2022>.
- Neukum, C., Azzam, R., 2009. Quantitative assessment of intrinsic groundwater vulnerability to contamination using numerical simulations. *Sci. Total Environ.* 408 (2), 245–254. <https://doi.org/10.1016/j.scitotenv.2009.09.046>.
- Pal, M., Makauskas, P., Malik, S., 2023. Upscaling porous media using neural networks: a deep learning approach to homogenization and averaging. *Processes* 11 (2), 601. <https://doi.org/10.3390/pr11020601>.
- Pan, L., Housh, M., Liu, P., et al., 2015. Robust stochastic optimization for reservoir operation. *Water Resour. Res.* 51 (1), 409–429. <https://doi.org/10.1002/2014WR015380>.
- Peaceman, D.W., 1993. Representation of a horizontal well in numerical reservoir simulation. *SPE Adv. Technol.* 1 (1), 7–16. <https://doi.org/10.2118/21217-PA>.
- Prabhune, B., Suresh, K., 2024. An isoparametric tangled finite element method for handling higher-order elements with negative Jacobian. *Comput. Mech.* 73 (1), 159–176. <https://doi.org/10.1007/s00466-023-02361-4>.
- Qi, D., Hesketh, T., 2005. An analysis of upscaling techniques for reservoir simulation. *Petrol. Sci. Technol.* 23 (7–8), 827–842. <https://doi.org/10.1081/LFT-200033132>.
- Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>.
- Ramos, E.M., Borges, M.R., Giraldo, G.A., et al., 2023. Prediction of permeability of porous media using optimized convolutional neural networks. *Comput. Geosci.* 27 (1), 1–34. <https://doi.org/10.1007/s10596-022-10177-z>.
- Rezaei, S., Asl, R.N., Faroughi, S., et al., 2025. A finite operator learning technique for mapping the elastic properties of microstructures to their mechanical deformations. *Int. J. Numer. Methods Eng.* 126 (1), e7637. <https://doi.org/10.1002/nme.7637>.
- Rubin, Y., 2003. *Applied Stochastic Hydrogeology*. Oxford University Press, New York. <https://doi.org/10.1093/oso/9780195138047.001.0001>.
- Salehian, M., Sefat, M.H., Muradov, K., 2021. Robust integrated optimization of well placement and control under field production constraints. *J. Petrol. Sci. Eng.* 205, 108926. <https://doi.org/10.1016/j.petrol.2021.108926>.
- Santos, A., Scanavini, H.F.A., Pedrini, H., et al., 2022. An artificial intelligence method for improving upscaling in complex reservoirs. *J. Petrol. Sci. Eng.* 211, 110071. <https://doi.org/10.1016/j.petrol.2021.110071>.
- Sayyafzadeh, M., Guéillot, D., 2022. Rapid permeability upscaling using convolutional neural networks. In: *ECMOR 2022. European Association of Geoscientists & Engineers, The Hague, Netherlands/Online*, pp. 1–12. <https://doi.org/10.3997/2214-4609.202244018>.
- Siraj, M.M., Van den Hof, P.M.J., Jansen, J.D., 2016. Robust optimization of waterflooding in oil reservoirs using risk management tools. *IFAC-PapersOnLine* 49 (7), 133–138. <https://doi.org/10.1016/j.ifacol.2016.07.229>.
- Tang, M., Liu, Y., Durlflosky, L.J., 2020. A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. *J. Comput. Phys.* 413, 109456. <https://doi.org/10.1016/j.jcp.2020.109456>.
- Tariq, Z., Fu, Q., Abu-Al-Saud, M.O., et al., 2025. A comparative study of deep learning-based simulation for geological CO₂ sequestration. *Adv. Water Resour.* 205, 105096. <https://doi.org/10.1016/j.advwatres.2025.105096>.
- Tegnander, C., Gimse, T., 1998. Flow simulations to evaluate upscaling of permeability. *Math. Geol.* 30 (6), 717–731. <https://doi.org/10.1023/A:1022347403212>.
- Wang, C., Li, G., Reynolds, A.C., 2009. Production optimization in closed-loop reservoir management. *SPE J.* 14 (3), 506–523. <https://doi.org/10.2118/109805-PA>.
- Wang, N., Chang, H., Zhang, D., 2022. Surrogate and inverse modeling for two-phase flow in porous media via theory-guided convolutional neural network. *J. Comput. Phys.* 466, 111419. <https://doi.org/10.1016/j.jcp.2022.111419>.
- Wang, N., Liao, Q., Chang, H., et al., 2023. 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, N., Zhang, D., Chang, H., et al., 2020. Deep learning of subsurface flow via theory-guided neural network. *J. Hydrol.* 584, 124700. <https://doi.org/10.1016/j.jhydrol.2020.124700>.
- Wang, S., Wang, H., Perdikaris, P., 2021. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. *Sci. Adv.* 7 (40), eabi8605. <https://doi.org/10.1126/sciadv.abi8605>.
- Wang, Y.J., Jin, Y., Lin, B.T., et al., 2025. Deep learning-based upscaling for CO₂ injection into saline aquifers. *Pet. Sci.* 22 (4), 1712–1735. <https://doi.org/10.1016/j.petsci.2025.02.003>.
- Wen, G., Li, Z., Azizzadenesheli, K., et al., 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>.
- White, C.D., Horne, R.N., 1987. Computing absolute transmissibility in the presence of fine-scale heterogeneity. In: *SPE Symposium on Reservoir Simulation*, SPE-16011-MS. <https://doi.org/10.2118/16011-MS>.
- White, L., Brandman, J., Trenev, D., 2021. Consistency and local conservation in finite-element simulation of flow in porous media. *Comput. Geosci.* 25 (3), 1123–1138. <https://doi.org/10.1007/s10596-021-10044-3>.
- Wu, X.H., Efendiev, Y., Hou, T.Y., 2002. Analysis of upscaling absolute permeability. *Discrete and Continuous Dynamical Systems - B* 2 (2), 185–204. <https://doi.org/10.3934/dcdsb.2002.2.185>.
- Wu, Y.S., 2016. *Multiphase Fluid Flow in Porous and Fractured Reservoirs*. Elsevier, Amsterdam. <https://doi.org/10.1016/C2015-0-00766-3>.
- Xu, T., Liu, D., Hao, P., et al., 2024. Variational operator learning: a unified paradigm marrying training neural operators and solving partial differential equations. *J. Mech. Phys. Solid.* 190, 105714. <https://doi.org/10.1016/j.jmps.2024.105714>.
- Yan, X., Huang, Z., Yao, J., et al., 2016. An efficient embedded discrete fracture model based on mimetic finite difference method. *J. Petrol. Sci. Eng.* 145, 11–21. <https://doi.org/10.1016/j.petrol.2016.03.013>.
- Yan, X., Huang, Z., Yao, J., et al., 2019. Numerical simulation of hydro-mechanical coupling in fractured vuggy porous media using the equivalent continuum model and embedded discrete fracture model. *Adv. Water Resour.* 126, 137–154. <https://doi.org/10.1016/j.advwatres.2019.02.013>.
- Yan, X., Lin, J., Ju, Y., et al., 2025. A finite-volume based physics-informed Fourier neural operator network for parametric learning of subsurface flow. *Adv. Water Resour.* 205, 105087. <https://doi.org/10.1016/j.advwatres.2025.105087>.
- Yan, X., Lin, J., Wang, S., et al., 2024. Physics-informed neural network simulation of two-phase flow in heterogeneous and fractured porous media. *Adv. Water Resour.* 189, 104731. <https://doi.org/10.1016/j.advwatres.2024.104731>.
- Yang, G., Xu, R., Tian, Y., et al., 2024. Data-driven methods for flow and transport in porous media: a review. *Int. J. Heat Mass Tran.* 235, 126149. <https://doi.org/10.1016/j.jheatmasstransfer.2024.126149>.
- Yi, J., Li, H., Shan, X., et al., 2022. Division and identification of vertical reservoir units in Archeozoic metamorphic buried hill of Bozhong Sag, Bohai Bay Basin, East China. *Petrol. Explor. Dev.* 49 (6), 1282–1294. [https://doi.org/10.1016/S1876-3804\(23\)60349-0](https://doi.org/10.1016/S1876-3804(23)60349-0).
- Zhang, P., Pickup, G., Christie, M., 2008. A new practical method for upscaling in highly heterogeneous reservoir models. *SPE J.* 13 (1), 68–76. <https://doi.org/10.2118/103760-PA>.
- Zhang, Z., Yan, X., Liu, P., et al., 2023. A physics-informed convolutional neural network for the simulation and prediction of two-phase Darcy flows in heterogeneous porous media. *J. Comput. Phys.* 477, 111919. <https://doi.org/10.1016/j.jcp.2023.111919>.
- Zhang, Z., Agarwal, R., 2013. Numerical simulation and optimization of CO₂ sequestration in saline aquifers. *Comput. Fluid* 80, 79–87. <https://doi.org/10.1016/j.compfluid.2012.04.027>.
- Zhao, Q., Wang, Y., Chen, C., 2024. Numerical simulation of the impact of different cushion gases on underground hydrogen storage in aquifers based on an experimentally-benchmarked equation-of-state. *Int. J. Hydrogen Energy* 50, 495–511. <https://doi.org/10.1016/j.ijhydene.2023.07.262>.