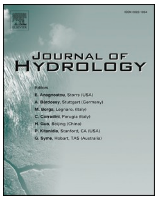
Journal of Hydrology 665 (2026) 134701

Contents lists available at ScienceDirect

Journal of Hydrology

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



Research papers

Combining hydrologic, chemical, and geophysical deep learning-based inversion for heterogeneous aquifer structure identification

Yuzhou Xia [a], Chuanjun Zhan [b,c,*], Zhenxue Dai [a,b,*], Jichun Wu [d], Xiaoying Zhang [a], Huichao Yin [e], Lin Zhu [f], Jiahe Yan [g], Zihao Wang [a], Mohamad Reza Soltanian [h], Kenneth C. Carroll [e]

[a] College of Construction Engineering, Jilin University, Changchun 130026, China
[b] School of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao 266520, China
[c] Department of Earth Sciences, The University of Hong Kong, Hong Kong
[d] Key Laboratory of Surficial Geochemistry of Ministry of Education, School of Earth Sciences and Engineering, Nanjing University, Nanjing, China
[e] Department of Plant & Environmental Sciences, New Mexico State University, Las Cruces, NM 88003, USA
[f] College of Resource Environment and Tourism, Capital Normal University, Beijing, China
[g] College of Geo-Exploration Science and Technology, Jilin University, Changchun 130026, China
[h] Departments of Geosciences and Environmental Engineering, University of Cincinnati, Cincinnati, OH, USA

ARTICLE INFO

This manuscript was handled by Renato Morbidelli, Editor-in-Chief

Keywords:
Heterogeneous aquifer structure
Inverse modeling
Electrical resistivity tomography
Data assimilation
Deep learning
Surrogate modeling

ABSTRACT

Accurate identification of heterogeneous aquifer structures is critical for geoscientific research. To address the limitations of inversion accuracy caused by sparse observations and the high computational cost of the inversion process, this study develops a Combined Hydrologic, Chemical, and Geophysical Data Deep Inversion Framework (HCGDIF). The framework integrates facies observations, hydraulic head, solute concentration, and electrical resistivity tomography (ERT) data, and achieves efficient generation of heterogeneous structures through the tight integration of hydrologic and geophysical models with deep learning (DL) techniques. The performance of the framework is evaluated through a synthetic ERT-monitored tracer experiment. Results demonstrate that the Integrated Structure Generation Model not only enables the parameterization of high-dimensional heterogeneous structures, but also generates a large number of random realizations consistent with prior geostatistical knowledge from limited facies observations, thereby providing reliable training data for DL models. The Hybrid Residual Octave-Convolution Network with Squeeze-and-Excitation Attention (HROCN) effectively captures the complex nonlinear mapping between aquifer structures and ERT data, making prediction approximately 33 times faster. Inversion results with different data types further indicate that, compared with direct estimation from facies observations alone, assimilating at least one observation type can improve identification accuracy. Among them, ERT data yields the best performance, increasing average identification accuracy by 13.9 % and reducing uncertainty by 4.31 %. Moreover, multi-source data fusion enhances the accuracy relative to single data, but also introduces additional uncertainty, while the inclusion of ERT data helps mitigate this uncertainty.

1. Introduction

Accurate identification of the heterogeneous aquifer structures remains a significant challenge in geoscientific research (Hart, 1999; Magalhães et al., 2020). Compared to inappropriate model parameters, uncertainty from aquifer structure often leads a more considerable misfit to the numerical simulations (Dai et al., 2020; Wallace and Soltanian, 2021). Therefore, identifying a reliable aquifer model is essential for groundwater remediation (Bianchi and Pedretti, 2018; Cirpka et al., 2022), water resources management (McGarr et al., 2021), hazardous

---

Abbreviations: DA, data assimilation; HCGDIF, Combined Hydrologic, Chemical, and Geophysical Data Deep Inversion Framework; ERT, Electrical Resistivity Tomography; DL, deep learning; DOCRN, Deep Octave Convolution Residual Network; HROCN, Hybrid Residual Octave Convolution Network with Squeeze-and-Excitation Attention; OctRRDRB, Octave Residual in Residual Dense Residual Block; OctRDB, Octave Residual Dense Block; CNB, ConvNeXt Block; DW, depth-wise convolution; ILUES, Iterative Local Updating Ensemble Smoother.

* Corresponding authors at: College of Construction Engineering, Jilin University, Changchun 130026, China.
E-mail addresses: zhanchuanjun@qut.edu.cn (C. Zhan), dzx@jlu.edu.cn (Z. Dai).

https://doi.org/10.1016/j.jhydrol.2025.134701
Received 18 April 2025; Received in revised form 26 November 2025; Accepted 28 November 2025

waste disposal (Jia et al., 2023; Ma et al., 2024), and geological storage of carbon dioxide and hydrogen (Xu et al., 2025).

The early mainstream method for identifying aquifer structures is the stochastic simulation methods, which establish relationships between facies at different spatial locations through spatial correlation functions and subsequently interpolate facies based on these relationships (Carle and Fogg, 1996). However, the effectiveness of stochastic simulation methods depends on the spatial coverage and data density of facies observations (e.g., drilling data) in the study area (Chen and Rubin, 2003; Soltanian and Ritzi, 2014). Further research has shown that dynamic observations, such as hydraulic heads (Berg and Illman, 2011; Cardiff et al., 2013; Xiang et al., 2009; Yeh and Liu, 2000; Zhu and Yeh, 2005) and solute concentrations (Mo et al., 2020; Xu and Gómez-Hernández, 2016) exhibit unique response characteristics to specific heterogeneous aquifer structures (Brunetti et al., 2019; Khambhammettu et al., 2020; Kim et al., 2018; Liu et al., 2019; Reuschen et al., 2021). Therefore, data assimilation (DA) methods based on dynamic observations can effectively address the reliance of stochastic simulation methods on facies observations (Bravo et al., 2002; Dausman et al., 2010; Song et al., 2019; Soueid Ahmed et al., 2014; Zhang et al., 2020). In theory, using more dynamic observations can improve the reliability of inversion results, but the methods for acquiring dynamic observations, such as tracer tests and pumping tests that involve injection/pumping wells and observation wells are essentially intrusive survey methods. On the one hand, survey costs are high; on the other hand, the structural damage caused by the sampling process can affect groundwater flow and solute transport, increasing the uncertainty of experimental results (Chen et al., 2022). Therefore, balancing the need for high-quality dynamic observations with cost-effective survey methods is a critical challenge for traditional aquifer structure identification.

Geophysical methods, with their non-intrusive nature and high-density sampling, provide cost-effective and abundant spatially continuous information, offering a valuable solution to address the limitations of traditional methods in identifying heterogeneous aquifer structures (Christensen et al., 2016). Among various geophysical methods, Electrical Resistivity Tomography (ERT) is particularly favored by hydrogeophysicists for its high sensitivity to variations in water content and solid phase (e.g., soil and aquifer mineralogy and grain size distributions) (Ikard et al., 2023; Mao et al., 2022; Mao et al., 2024; Rucker et al., 2021). DA methods incorporating geophysical data enable the coupled inversion of geophysical and hydrologic models (Hubbard et al., 2001; Yeh and Šimůnek, 2002). This approach eliminates the need for traditional geophysical inversion and instead focuses on mutual constraints between models to identify physical laws from raw geophysical monitoring data (Binley et al., 2015). Existing studies demonstrate the significant potential of geophysical data in joint assimilation with hydraulic head and concentration data (Pollock and Cirpka, 2010; Pollock and Cirpka, 2012; Yeh et al., 2008). Algorithms such as the Ensemble Kalman Filter (EnKF) and Ensemble Smoother with Multiple Data Assimilation (ESMDA) have successfully integrated various data types in sandbox experiments and regional large-scale studies (Camporese et al., 2015; Kang et al., 2018; Kang et al., 2019; Neven and Renard, 2023; Xu et al., 2021). However, DA methods face challenges in high-dimensional parameter inversion during application, and obtaining responses requires repeatedly running forward models, especially for complex and large geophysical dynamic responses, which demand substantial computational time.

The rapid development of machine intelligence, especially deep learning (DL), with its superior data compatibility and ability to characterize nonlinear relationships, has been used to develop structure generation models and surrogate models for responses in DA frameworks, thereby elevating aquifer structure identification methods to an unprecedented new stage (Liu et al., 2019; Song et al., 2019; Wang et al., 2021). For structure generation models, generative adversarial networks (GANs) (Laloy et al., 2018; Zhang et al., 2019) and variational autoencoders (VAEs) (Canchumuni et al., 2017; Laloy et al., 2017) have been effectively applied. Later research enhanced these models by integrating advanced models (Chan and Elsheikh, 2019; Chen, 2024; Deng, 2024; Han et al., 2022; Jamil, 2024; Lopez-Alvis et al., 2022; Mo et al., 2020). Additionally, DL-based surrogate models for forward simulation can speed up the process of acquiring dynamic responses (Siade et al., 2020; Xiao et al., 2021). Several classical models from the field of computer vision can be applied to surrogate modeling, such as ResNet and DenseNet. However, DA methods aim to extract local heterogeneity information of aquifer structures from geophysical raw observations. Therefore, models should be designed to map hydrologic parameters to geophysical raw observations. Currently, no studies have utilized models to predict geophysical raw observations within the framework of DA methods.

This study develops a Combined Hydrologic, Chemical, and Geophysical Data Deep Inversion Framework (HCGDIF) for identifying heterogeneous aquifer structures. The datasets required to train all DL models within the framework are generated through stochastic simulations and high-fidelity numerical simulations. A synthetic case study is conducted to evaluate the performance of the structure generation model and surrogate models embedded in the framework. Based on this setup, multiple inversion scenarios are designed using different types of data, including facies observations, hydraulic head, solute concentration, and ERT data. The performance of HCGDIF is assessed by analyzing the accuracy and uncertainty of the inversion results.

The paper is organized as follows. Section 2 outlines the methods, including the structure of the proposed framework, integrated structure generation model, the coupled hydrogeophysical forward model, DL-based surrogate models, and DA algorithm. Section 3 presents the designed application example, including the real structure and cases with different data combinations. Section 4 discusses the identification results of various cases. Section 5 summarizes the conclusions of this work.

2. Methods

2.1. Combined hydrologic, chemical, and geophysical data deep inversion framework

The HCGDIF has four modules: (1) Integrated Structure Generation Model (Fig. 1a, Fig. 1c); (2) coupled hydrogeophysical forward model (Fig. 1b); (3) DL-based surrogate models (Fig. 1b); (4) DA algorithm. Sections 2.2-2.5 provide a comprehensive account of the four modules. The Integrated Structure Generation Model is designed to parameterize aquifer structures, employing low-dimensional latent vectors to represent aquifer structures during parameter inversion, thereby alleviating the computational burden associated with high-dimensional parameter inversion. The coupled hydrogeophysical forward model is utilized to construct a dataset linking aquifer structures to multi-source dynamic responses. The trained surrogate models effectively capture the highly nonlinear relationship between aquifer structures (inputs) and dynamic responses (outputs), and thus significantly reducing the computational time required for high-fidelity forward simulations.

The inversion workflow is illustrated in Fig. 1d. First, a set of prior latent vectors is randomly sampled from a standard normal distribution and input into the structure generation model to obtain a prior structure ensemble. Subsequently, the surrogate models are employed to simulate multi-source dynamic responses for the prior structures, generating the spatiotemporal distributions of solute concentration, hydraulic head, and ERT data that correspond to specific heterogeneous facies configurations. Finally, the DA algorithm updates the prior latent vectors by minimizing the discrepancies between dynamic observations and responses, enabling precise adjustments to aquifer structures and yielding posterior aquifer structures with high accuracy.

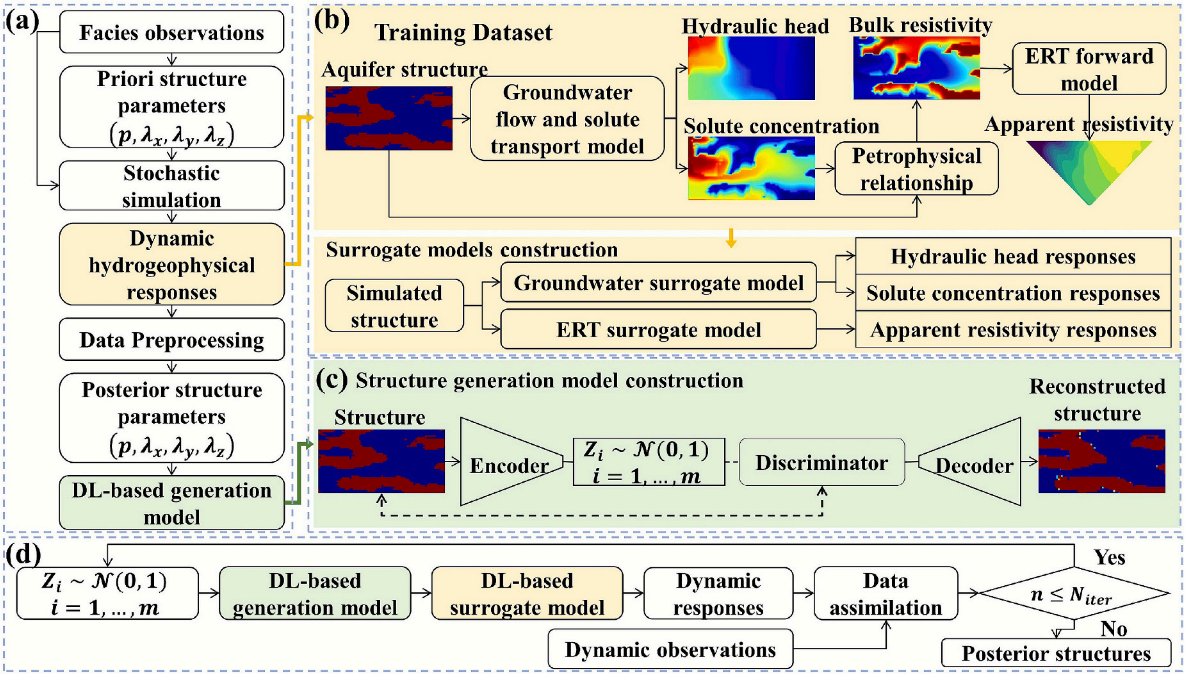


Fig. 1. Flow diagram of HCGDIF inversion process. (a) Stochastic simulation-based structure generation model, (b) Coupled hydrogeophysical forward model and surrogate models, (c) DL-based structure generation model, (d) Inversion of aquifer structure based on DA algorithms and multi-source dynamic observations.

2.2. *Integrated structure generation model*

The Integrated Structure Generation Model consists of two sequential sub-models: the stochastic simulation-based structure generation model (Fig. 1a), which incorporates facies observations, and the DL-based structure generation model (Fig. 1c), which parameterizes high-dimensional aquifer structures. In the stochastic simulation-based structure generation model, a broad geostatistical parameter distribution range (volume proportions p and mean lengths L) is initially established based on site-specific facies observations. Structure parameters are then randomly sampled from this range and used for stochastic simulation with a transition probability model. The analytical solution of the transition probability model is employed to represent facies spatial correlations (Dai et al., 2007):

$$T_{ij}(\sigma_\phi) = (\delta_{ij} - p_i) e^{-\frac{\sigma_\phi}{L_i^\phi (1-p_i)}} + p_j \tag{1}$$

where $T_{ij}(\sigma_\phi)$ represents the transition probability from facies i to facies j at lag distance σ_ϕ in direction ϕ, δ_{ij} is the Kronecker delta, p_i and p_j are the proportions of facies i and facies j, L_i^ϕ is the mean length of facies i in direction ϕ. Subsequently, indicator cokriging simulation is applied to complete the stochastic simulation of aquifer structures (Carle and Fogg, 1996). The stochastic modeling software GEOST (Dai et al., 2014) is used to conduct the aforementioned analyses and computations.

The Octave Convolution Adversarial Autoencoder (Zhan et al., 2021) is applied in the DL-based structure generation model to address the limitations of conventional GAN-based generators. Although GANs can capture complex data distributions, they often suffer from mode collapse, where different latent inputs produce nearly identical aquifer structures. This reduces the structural diversity and limits the flexibility of subsequent data fusion processes. The Octave Convolution Adversarial Autoencoder introduces adversarial learning into a variational autoencoder framework, combining the stable latent representation of variational autoencoder with the realism enhancement of adversarial training. This design introduces adversarial perturbations that enhance the model's robustness and generalization, allowing it to generate structures that more closely match the spatial heterogeneity present in the training data, which cannot be achieved by conventional GANs. The training dataset required for the generation model is generated using stochastic structures established by GEOST. However, the broad prior parameter ranges determined through manual experience often result in high uncertainty and limited accuracy in the generated stochastic structures, deviating significantly from real structures. This quality of datasets is suboptimal for training the generation model, necessitating preprocessing of the prior parameter set using a DA algorithm. Structure parameters obtained after inversion, combined with facies observations, are input into GEOST to generate high-quality stochastic simulation structure datasets. This process enables the generation model to focus on learning features resembling actual aquifer structures. Detailed information on data preprocessing and the hyperparameters of the DL-based structure generation model is provided in Supporting Information Text S1.

As shown in Fig. 1c, the encoder encodes the stochastic structure into a latent vector following a standard normal distribution. The discriminator evaluates whether the generated latent vector originates from the training set, while the decoder decodes the latent vector into a reconstructed structure with features similar to the stochastic structure. After training, the network model and weight files of the decoder are saved, enabling the transformation of low-dimensional stochastic latent vectors into high-dimensional aquifer structures. Consequently, the DA algorithm can refine aquifer structures by optimizing the latent vectors. Theoretically, a larger latent vector dimensionality improves the model's ability to adjust aquifer structures in detail; however, it also significantly increases the computational burden of parameter inversion. Thus, the latent vector size should be tailored to specific tasks to balance computational performance and cost.

2.3. *Coupled hydrogeophysical forward model*

2.3.1. *Groundwater flow and solute transport model*

The governing equation for the steady-state flow field in this study is (Bear, 1979):

$$\nabla K \nabla H = 0 \tag{2}$$

where K and H represent the hydraulic conductivity (m/d) and hydraulic head (m) respectively. The boundary conditions are considered as:

$$\begin{cases} H(x,y,t)|\Gamma_{left} = H_{left} \\ H(x,y,t)|\Gamma_{right} = H_{right} \\ K\dfrac{\partial H}{\partial n}|\Gamma_{top} = 0 \\ K\dfrac{\partial H}{\partial n}|\Gamma_{bottom} = 0 \end{cases} \tag{3}$$

in which H_{left} and H_{right} are prescribed head at the left boundary Γ_{left} and right boundary Γ_{right}. The top boundary Γ_{top} and bottom boundary Γ_{bottom} are impervious.

The governing equation for nonreactive contaminant transport is:

$$\frac{\partial c}{\partial t} = \nabla \bullet (\boldsymbol{D} \bullet \nabla(c)) - \nabla \bullet (\boldsymbol{u}c) + S \tag{4}$$

where c is the solute concentration (mol/m^3), t is the time (s), $\boldsymbol{D}$ is the diffusion coefficient tensor (m^2/s), $\boldsymbol{u}$ is the pore water velocity (m/s), and S is the sink/source (mol/m^3 • s). The boundary conditions are considered as:

$$\begin{cases} c(x,y,t)|\Gamma_{left} = c_0 \\ c(x,y,t)|\Gamma_{right} = c_0 \\ c(x,y,t)|\Gamma_{top} = c_0 \\ c(x,y,t)|\Gamma_{bottom} = c_0 \end{cases} \tag{5}$$

where c_0 presents the initial solute concentration. The groundwater flow and solute transport equations are solved by TOUGHREACT (Xu et al., 2006).

2.3.2. ERT forward model

Given that traditional geophysical inversion struggles to characterize the highly heterogeneous features of aquifer structures, HCGDIF utilizes apparent resistivity in the inversion process. The objective of the ERT forward model is to measure apparent resistivity data that characteristically responds to the heterogeneity of aquifer structures. The resistivity of water-saturated porous media is primarily governed by the resistivity of pore fluids and surface resistivity. In the aquifer model discussed in this paper, the influence of surface resistivity is neglected due to the absence of clay content (Camporese et al., 2015; Pollock and Cirpka, 2010). Archie's law is applied to describe the relationship between pore fluid resistivity ρ_w (Ω • m) and bulk resistivity ρ (Ω • m) (Han et al., 2024):

$$\rho = \phi^{-m} S_w^{-n} \rho_w \tag{6}$$

where ϕ represents porosity, m is the cementation exponent, n is the saturation exponent, and S_w is the water saturation. These parameters are site-specific empirical values, and applying these equations to other locations would require recalibration of the associated parameters. ρ_w is influenced by both temperature T (℃) and c (mol/m^3) (Sen and Goode, 1992):

$$\frac{1}{\rho_w} = \left(5.6 + 0.27T - 1.5 \times 10^{-4} T^2\right) \bullet c - \frac{2.36 + 0.099T}{1 + 0.214c} \bullet c^{\frac{3}{2}} \tag{7}$$

According to equations (6) and (7), under water-saturated conditions, the aquifer bulk resistivity ρ is primarily influenced by ϕ, which has a static heterogeneous distribution, and c, which has a dynamic distribution. Therefore, the apparent resistivity data distribution in this study exhibits dual variability in both temporal and spatial dimensions, meaning that this time-lapse ERT measurement can offer more valuable information for identifying subsurface sedimentary structures (Jardani et al., 2013; Kang et al., 2019).

In this study, the ERT simulation utilizes a point current electrode to inject direct current, generating a current I (A) at the position (x_0, z_0). After determining the spatial distribution of $\rho(\Omega \bullet \text{m})$ the target potential is computed by solving the following equation (LaBrecque et al., 1996):

$$\nabla \cdot \left(\frac{1}{\rho} \nabla \widetilde{V}_k\right) - k^2 \frac{1}{\rho} \widetilde{V}_k = -I\delta(x - x_0)\delta(z - z_0) \tag{8}$$

where $\widetilde{V}_k(x, z)$ (V) is potential in the Fourier transform domain of wavenumber k and δ denotes the Dirac function. The boundary conditions are considered as:

$$\begin{cases} \dfrac{\partial \widetilde{V}_k}{\partial n} + \widetilde{V}_k k \cos\theta \dfrac{K_1(kr)}{K_0(kr)} = 0 \in \Gamma_s \\ \dfrac{\partial \widetilde{V}_k}{\partial n} = 0 \in \Gamma_\infty \end{cases} \tag{9}$$

where Γ_s is surface boundary and $\boldsymbol{r}$ is the distance between the supply electrode and Γ_s; K_0 and K_1 respectively denote the zero-order and first-order modified Bessel function of the second kind; and Γ_∞ denotes the truncated boundary. Equations 8–9 are solved using SimPEG (Cockett et al., 2015; Heagy et al., 2017).

2.4. DL-based surrogate models

To obtain the multisource response state fields of reconstructed structures in the inversion process, the HCGDIF organizes the high-precision data obtained in Section 2.3 into corresponding datasets to train DL-based surrogate models, thereby accelerating the prediction process of multisource dynamic responses. For hydrological responses such as solute concentration and hydraulic head, we adopt the Deep Octave Convolution Residual Network (DOCRN), which is well-suited for image-to-image mapping tasks such as predicting spatial distributions of state variables from structure inputs (Zhan, 2022). Unlike hydrological responses, ERT data are vectorized and directly incorporated into the HCGDIF as one-dimensional inputs. This distinct data representation, combined with the more complex mapping from structure to ERT data, makes it unsuitable to simply extend the output channels of DOCRN. Such extension could significantly degrade prediction accuracy. Therefore, in this study, we design a Hybrid Residual Octave Convolution Network with Squeeze-and-Excitation Attention (HROCN) specifically for predicting ERT data. The errors generated in the inversion process primarily stem from dynamic observations and the surrogate model's prediction of dynamic responses. Errors in dynamic observations depend on manual techniques and the inherent accuracy of sensor equipment, making them unavoidable and beyond the scope of this study. To enhance the inversion effectiveness of the proposed framework, the surrogate model should possess accurate image-translation capabilities, which depend on the ability of the neural network to extract image features. Therefore, we integrated advanced feature extraction blocks into surrogate models to enhance the prediction accuracy of the responses.

Octave Residual in Residual Dense Residual Block (OctRRDRB, Fig. 2b) combines two key techniques: octave convolution and a deeply nested residual structure (Rakotonirina and Rasoanaivo, 2020). Octave convolution separates input features into low- and high-frequency components (Chen et al., 2019). This allows the model to focus more on high-frequency details, such as sharp facies boundaries, making it especially effective for initial feature extraction. The basic structure of the Octave Residual Dense Block (OctRDB, Fig. 2a) includes several convolutional groups. Every two convolutional groups are connected through residual connections, while the first four groups are also connected through dense connections. The deeply nested residual structure includes three residual dense blocks connected with skip connections

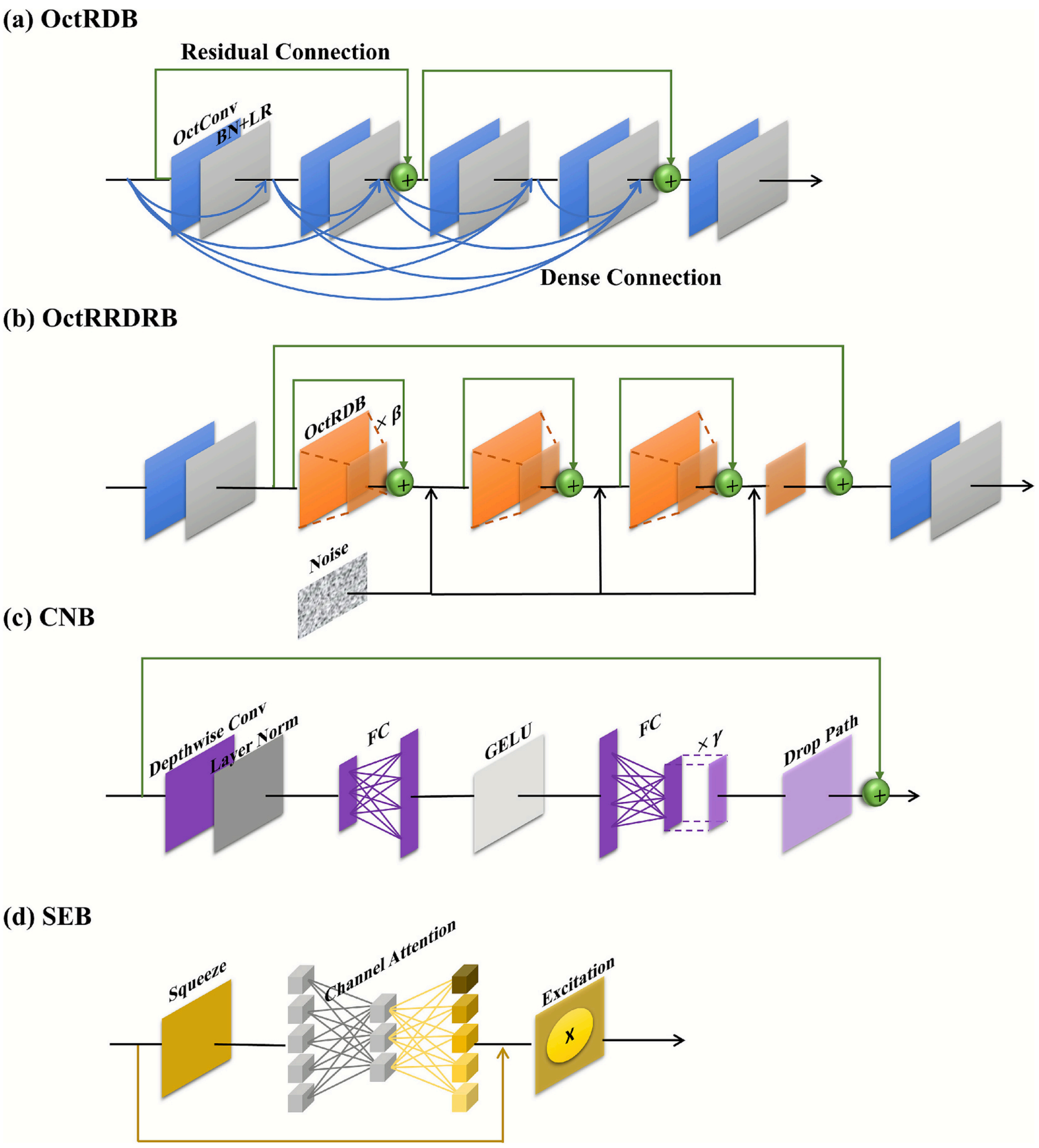


Fig. 2. The architecture of (a) OctRDB, (b) OctRRDRB, (c) CNB, (d) SEB. BN + LR means that the feature maps are processed sequentially through a Batch Normalization operation and then a LeakyRelu activation.

scaled by a factor β (here, β=0.2 (Szegedy et al., 2017)). This residual architecture enables the deep learning model to extract structural features of the input image across different scales. In the context of aquifer structure identification, such a design allows the model to capture facies distribution characteristics at multiple spatial scales, thereby improving its ability to represent complex heterogeneity. To reduce overfitting, Gaussian noise is added to the scaled outputs, as commonly done in generative tasks (Karras et al., 2021). Finally, all standard convolutions in the block are replaced with octave convolutions, forming the final OctRRDRB.

ConvNeXt Block (CNB) is the fundamental building block of ConvNext (Liu et al., 2022). ConvNeXt was designed to explore the performance limits of traditional convolutional networks against the backdrop of Transformers, which is dominating the field of visual recognition. ConvNeXt has demonstrated that its inference speed and accuracy have surpassed those of Swin Transformer (Liu et al., 2021) in a variety of vision tasks, including ImageNet classification (Deng et al., 2009), object detection/segmentation on COCO (Lin et al., 2014), and semantic segmentation on ADE20K (Zhou et al., 2019). The structure of CNB is shown in Fig. 2c, which adopts a more aggressive depthwise convolution at the module's start, inspired by the grouped convolution in ResNext (Zhou et al., 2019). Compared with standard convolution, depthwise convolution significantly reduces the number of parameters and computational cost by separating spatial and channel-wise feature extraction. This design allows the model to focus more efficiently on the spatial variations of facies features while maintaining rich channel information, thus improving computational efficiency and feature representation capability. In depthwise convolution, each convolution kernel has a channel size of 1, requiring the number of kernels to equal the number of channels in the input feature matrix. The depthwise convolution process can be described as:

$$Y^{1*1*C} = \sum_{i=1}^{C} X^{L_H*L_W*i} * V^{V_s*V_s*1} \tag{10}$$

where C represents the number of channels of the feature map, L_H and L_W represent the spatial dimensions of feature map X, and V_s represents the size of the convolution kernel V (here, K=7 (Liu et al., 2022)). For each input channel, the convolution operation outputs a compressed feature map Y with the same number of channels. Then, two fully-connected layers are used to expand and compress the features by a factor of 4, ensuring that the input and output feature map dimensions remain unchanged. Finally, the normalized output features from the fully-connected layer are connected residually to the feature map before the input depthwise convolution layer:

$$\widetilde{X}^{L_H{}^*L_W{}^*C} = X^{L_H{}^*L_W{}^*C} + \mathscr{D}\left(\gamma F_2\left(gF_1\left(\mathscr{L}\left(Y^{1^*1^*C}\right)\right)\right)\right) \tag{11}$$

where $\mathscr{L}$ represents Layer Normalization to avoid extreme values in $Y.F_1$ and F_2 represent the first expanding fully-connected layer and the second compressing fully-connected layer, respectively, g represents the activation function GELU, γ is a learnable dynamic parameter (set initially as$\gamma = 1\text{e-}6$ in this study (Liu et al., 2022)). $\mathscr{D}$ represents the Droppath regularization method, which helps prevent network overfitting.

There are three subtle adjustments in CNB compared to OctRRDRB: in terms of normalization layers, CNB replaces Batch Normalization with Leaky Relu and reduces the frequency of normalization layers to only following depthwise convolution, a strategy that has proven effective in comparison to the Swin Transformer. In terms of activation functions, CNB draws from Transformer practices by replacing Leaky Relu with GELU and similarly eliminates the traditional convolution network's practice of using an activation function after each convolution layer, employing it only once between the two fully-connected layers. Regarding feature scaling, CNB no longer uses a fixed constant for scaling but employs a dynamic parameter γ, which is gradually adjusted through backpropagation during training to optimize network performance. This scaling mechanism in CNB helps stabilize the training process of deep networks, avoiding gradient explosion or vanishing phenomena, and significantly reduces computation and parameter count.

Squeeze Excitation Block (SEB) is inspired by Squeeze-and-Excitation Networks (Hu et al., 2018), is a channel attention mechanism that automatically learns the relative importance of each channel feature within a feature map. In this context, a channel refers to an individual feature map in the output of a convolutional layer, where each channel represents a specific type of learned feature. The SEB basic network structure is shown in Fig. 2d. First, global average pooling is used to squeeze the spatial information of the feature map, averaging each feature map into a single real value with a global receptive field. This process can be described as:

$$z_c = \frac{1}{L_H \times L_W} \sum\nolimits_{i=1}^{L_H} \sum\nolimits_{j=1}^{L_W} u_c(i,j) \tag{12}$$

where L_H and L_W represent the spatial dimensions of the feature map, u_c represents the c th feature map, and z_c represents the real value of the squeezed c th feature map.

Next, a set of FC layers is used to learn the channel features. These learned channel features represent the importance of each channel in the context of the task at hand. Specifically, the FC layers capture dependencies between channels and help identify which channels contribute more to the network's decision-making process. The learned channel features are then used to excite the initial feature map by scaling each channel according to its learned importance. This process can be described as:

$$\begin{cases} s_c = \sigma(W_2\delta(W_1 z_c)) \\ \widetilde{x}_c = s_c u_c \end{cases} \tag{13}$$

where σ and δ represent the Sigmoid and ReLU activation functions, respectively. W_1 and W_2 are the weight matrices for the fully-connected layers that control the expansion and compression of the channel information, with a scaling factor$\alpha = 16$ (Hu et al., 2018). s_c represents the learned channel attention, which is a scalar that indicates the relative importance of the c th feature map. The final output $\widetilde{x}_c$ is the excited version of the c th feature map, where the important channels are enhanced, and the less important ones are suppressed. In summary, channel feature refers to the inherent characteristics of each feature map in the convolutional layers, and channel attention refers to the learned importance of these channels, which helps the network focus on the most relevant features while suppressing less informative ones. This mechanism not only reduces the computational load but also improves the model's ability to generalize by emphasizing key features and suppressing noise. DOCRN and HROCN are built on the aforementioned blocks (Fig. 3) and they serve as the surrogate models for multi-source dynamic responses. The application of HROCN in this study can be represented as:

$$\begin{cases} \widetilde{X}_{D\times L_W\times L_H} = NN_{DOCRN}(X_{L_W\times L_H}) \\ \widetilde{X}_L = NN_{HROCN-SE}(X_{L_W\times L_H}) \end{cases} \tag{14}$$

where $X_{L_W\times L_H}$ represents the heterogeneous aquifer structure input with width L_W and height L_H; $\widetilde{X}_L$ represents the ERT data with L data points;$\widetilde{X}_{D\times L_W\times L_H}$ represents hydrologic dynamic response fields with D types; NN represents the corresponding neural network architecture.

Further details on the network configurations and the hyperparameters of DOCRN and HROCN are provided in Supporting Information Text S2. Unlike the hourglass structure of DOCRN, HROCN adopts a progressive design where the feature map size is gradually reduced while the number of channels increases as the network deepens. Specifically, the OctRRDRB is positioned at the model's initial stage, utilizing its ability to distinguish high-frequency and low-frequency information to enhance the features in areas of intense aquifer structure variation. After being processed twice consecutively through the OctRRDRB, the feature map is linked with the initial features (which are the feature representations derived directly from the input data or from the output of earlier layers) via a residual connection. This allows the network to retain and refine important information from the input. The feature map is then passed through four CNBs with gradually increasing channels to efficiently capture essential sample features at a low computational cost. Compression of the feature map is achieved via convolutional layers. At the end of the final CNB, the feature map reaches its maximum number of channels. At this point, the SEB's powerful channel attention mechanism is used to weight each feature channel, which is then multiplied by the input feature map to obtain the weighted feature map. Since in this study the ERT data are reshaped into a 1D vector to facilitate the direct connection between the HROCN output and the DA algorithm, the model's final layer is designed with FC layers to output data in the required dimensions. Notably, the HROCN can be flexibly adapted in its application; depending on specific task requirements, the model's final layer can be modified to a combination of convolution and upsampling layers to output image data.

2.5. Data assimilation algorithm

The fundamental logic of applying indirect observations in DA methods is:$D_{obs} = \widetilde{D}_{res} + e$, the error e between dynamic responses. $\widetilde{D}_{res}$ and dynamic observations D_{obs} arises mainly from the estimation bias of aquifer structures. Thus, DA methods can generate simulated structures closer to reality by minimizing the e. This study employs the Iterative Local Update Ensemble Smoother (ILUES, (Zhang et al., 2018)) as the DA algorithm for parameter estimation. The stability of this method has been validated in recent studies (Xia et al., 2023). The updated process of ILUES can be expressed as:

$$r_n^i = r_n^{i-1} + \frac{D + e_n^{i-1}\sqrt{N} - F\left(r_n^{i-1}\right)}{C_{MF}^{i-1}\left(C_{FF}^{i-1} + NC_D\right)} \tag{15}$$

where r_n^{i-1} and $r_n^i \in R = [r_1, r_2 \cdots r_m]$ represent the n th parameter value in the ensemble before and after the i th update, respectively. D denotes the

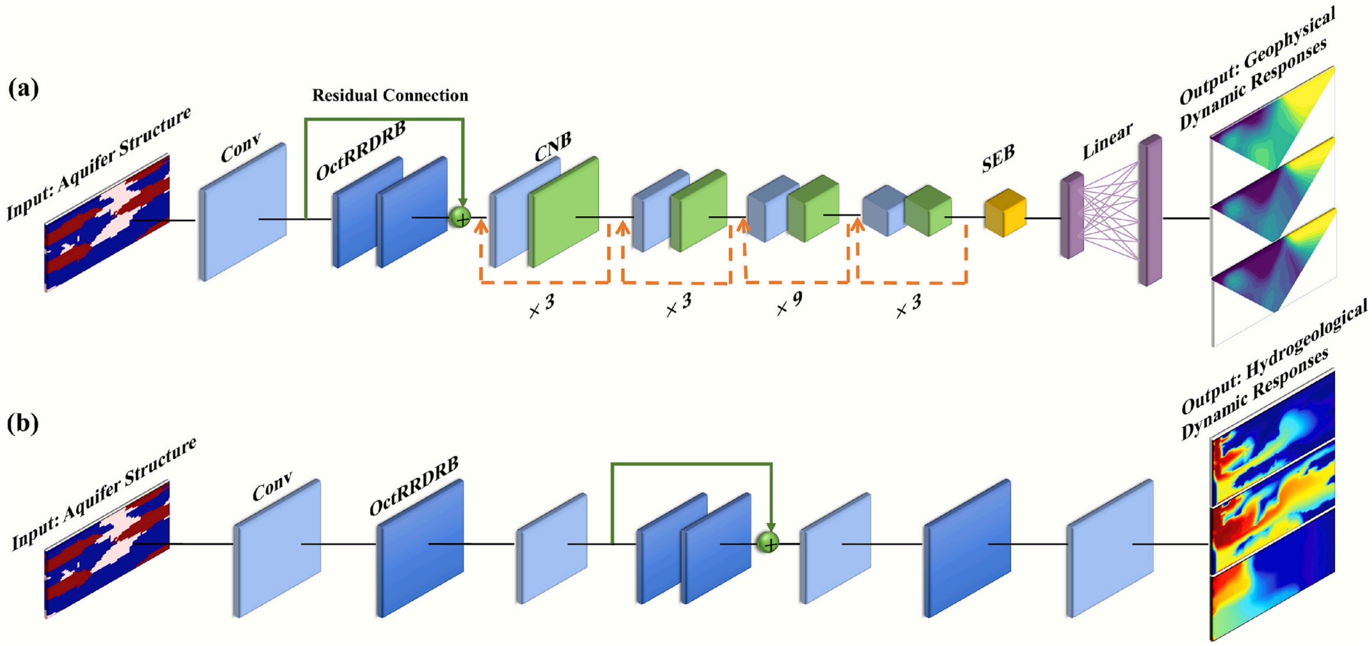


Fig. 3. The architecture of Surrogate Models: (a) HROCN, (b) DOCRN. The green solid line represents the residual connection, and the orange dashed line indicates the repeated input to the CNB. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

dynamic observations, N is the total number of iterations, e_n^{i-1} is the n th observation error, F represents the simulation process of the dynamic responses, C_D is the covariance of the observation errors. C_{MF}^{i-1} represents the cross-covariance between the parameters to be updated and the dynamic responses, and C_{FF}^{i-1} represents the auto-covariance of the dynamic responses.

3. Application example

A two-dimensional dual-facies heterogeneous aquifer was designed as the target for identification to evaluate the performance of HCGDIF. The model dimensions are 80 m × 20 m, discretized into 3200 grids (80 × 40), with each grid measuring 1 m × 0.5 m. Two facies types were incorporated into the model to represent high- and low-permeability zones, as illustrated in Fig. 4a. This dual-facies structure has been demonstrated to control flow and transport processes in subsurface environments (Gershenzon et al., 2015). The model was stochastically generated using GEOST, with a borehole placed every 20 m as conditional data. The corresponding true structure parameters (volume proportion p and mean lengths L) are shown in Table 1. Notably, the low-permeability facies has a smaller true volume proportion compared to high-permeability facies. When estimating the dual-facies structure using GEOST, each grid is initially assigned the facies type with the larger volume proportion. The spatial positions of the facies with the smaller volume proportion are then determined using the transition probability model, and the facies type of the corresponding grids is replaced. Due to the sparse spatial distribution of borehole data, the mean length in x direction L_x could not be accurately estimated, leading to a wider prior interval for L_x. In contrast, smaller prior intervals were assigned to the p and the mean length in z direction L_z. Note that all random sampling processes in this study utilized Latin hypercube

Table 1
Parameter Settings for Petrophysical Relationships.

Parameter	High-Permeability	Low-Permeability
Volume proportion (true / prior distribution)		0.4 / [0.35, 0.45]
Mean length in X-direction (true / prior distribution)		20 / [10, 80]
Mean length in Z-direction (true / prior distribution)		3 / [2.5, 3.5]
Permeability (10^{-10} m^2)	16.799	0.094
Porosity	0.29	0.28
Cementation exponentTemperature	2.6	2.6
(°C)	25	
Water resistivity (Ω·m)	20	

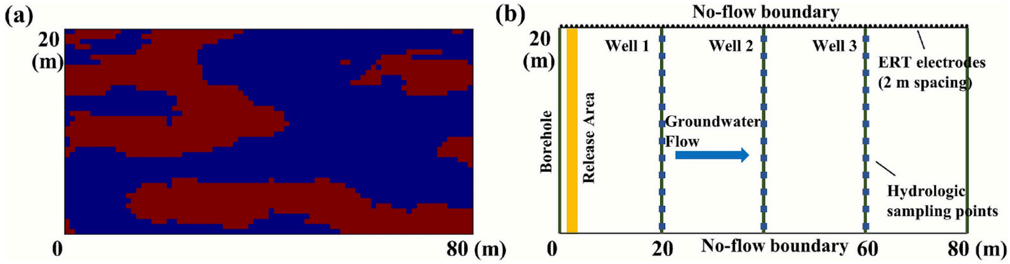


Fig. 4. (a) Distribution of low-permeability facies (red) and high-permeability facies (blue). (b) Sketch of the synthetic model domain. The green solid lines indicate the locations of facies observations, yellow area represents the release source zone, blue rectangles denote hydrologic sampling points, and black circles represent the electrodes distribution for the ERT survey along the land surface and upper boundary. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sampling, which provides more uniform sample coverage compared to Monte Carlo random sampling (McKay et al., 2000).

A synthetic salt water injection experiment is conducted using the model, with hydrologic sampling and ERT surveys employed to monitor solute transport. For the groundwater flow and solute transport model, the left and right sides are set as constant head boundaries with a hydraulic gradient of 5 × 10^{-3}, while other boundaries are no-flow boundaries. The selected hydraulic gradient represents a relatively small natural gradient typically observed in alluvial aquifers and is suitable for simulating the dispersion and migration of contaminant plumes (Chen et al., 2025). In contrast, larger hydraulic gradients, such as 0.05 and 0.1, are commonly found in large scale sandbox experiments or in situ test sites where flow conditions are enhanced by pumping wells (Jose et al., 2004; Zhang et al., 2015). We adopted a smaller natural gradient primarily to maintain the physical plausibility of the flow boundary conditions while ensuring that the velocity scale of the flow field approaches natural conditions. This allows the dominant control of stratigraphic heterogeneity on the spatiotemporal evolution of the solute plume to be revealed more clearly. An injection well is set at the left boundary as the injected-solute release source zone. The total simulation time is 150 days (after start of injection), with the injection well continuously releasing a 51.3 mol/m^3 NaCl solution at a rate of 0.01 kg/s for the first 15 days. The initial solute concentration of the entire region is set to 4.28 mol/m^3. To monitor the solute transport process, 12 randomly distributed concentration and pressure sampling points are installed in each of the three boreholes located in the middle of the model, for a total of 36 sampling points (Fig. 4b). To better capture dynamic response variations caused by local heterogeneity in the aquifer structure, the randomly selected sampling points should avoid low-permeability facies as much as possible. During the release period, concentration data and ERT data at six time points ($t = 2, 3, 4, 5, 7, 13$ days) and post-release head data ($t = 16$ days) are collected, resulting in 252 hydrologic data points for inversion, with 1 % observation noise considered (Zhanet al., 2022). For the ERT forward model, the relevant parameters in the petrophysical relationships are listed in Table 1. To maintain the comparability of the synthetic experiment and the simplicity of the model, each facies is treated as having representative constant parameters, porosity, cementation exponent, and saturation exponent, which are assumed invariant within the same facies. The selected parameters are derived from representative values extracted from actual measurements in the Chaobai River alluvial fan of the Beijing Plain (Zhu et al., 2017). In this study, two media with markedly different permeabilities were chosen for comparative filling, aiming to highlight as clearly as possible the controlling effect of structural heterogeneity on flow and transport. Furthermore, consistent with the treatment commonly adopted in synthetic case studies and sandbox experiments, we set $S_w = 1$ in Eq. (6), i.e., considering only the direct correspondence between bulk conductivity under saturated conditions and pore-water conductivity. This highlights how variations in solute concentration influence bulk conductivity through changes in pore-water conductivity. Electrodes are arranged on the top horizontal surface of the model at intervals of 2 m (Fig. 4b), employing the commonly used dipole–dipole array for field surveys (Johnson et al., 2015). A total of 1560 ERT data points are used for inversion, with 10 % observation noise included (Kang et al., 2021). This study designed five cases to represent different inversion scenarios, with detailed information provided in Table 2. The prior scenario directly identifies aquifer structures using stochastic simulation method based on facies observations, while cases 1–5 use the HCGDIF to identify aquifer structures by assimilating different types of dynamic observations. All inversion schemes share the same set of facies observations from boreholes as conditional data.

Table 2
Definition of identification scenarios for measurement types.

Case	Observations type	Number of data
Prior	F	200
1	$F + H$	236
2	$F + C$	416
3	$F + C + H$	452
4	F + ERT	1760
5	$F + C + H$ + ERT	2012

Abbreviations: *F*, facies observations from drilling; *C*, solute concentration; *H*, hydraulic head; ERT, electrical resistivity tomography.

4. Results and discussion

4.1. Training results of the structure generation model and surrogate models

4.1.1. Performance evaluation of integrated structure generation model

According to the method described in Section 2.2, GEOST are employed to generate 10,000 random structures from facies observations, forming the training set for the DL-based generation model. The training epochs are set to 50 and the batch size is 64. The learning rate is 0.002, and the adaptive optimization algorithm Adam is used to update neural network parameters based on gradient information. The dimensions of the latent vector are set to 200 to represent the high-dimensional aquifer structure (3200 dimensions), compressing the parameters to be estimated by 16 times, which significantly reduces the computational burden of parameter inversion. Fig. 5a-c show three aquifer structures stochastically generated by GEOST, with the corresponding reconstruction results shown in Fig. 5d-f. The visualized analysis indicates that the reconstruction results exhibit only slight estimation errors at the boundaries between different facies, with accuracies of 97.66 %, 97.38 %, and 97.25 % calculated grid by grid among the 3200 grids. Therefore, DL-based generation model can accurately capture the facies distribution patterns in the dataset, supporting the DA algorithm in refining aquifer structure adjustments.

4.1.2. Performance evaluation of DOCRN

DOCRN and HROCN are used as surrogate models for forward simulations within the proposed framework. These surrogate models need to ensure prediction accuracy while improving simulation speed. To obtain the dataset required for training, we randomly sample 10,000 latent vectors from a standard normal distribution and input them into the decoder of the DL generation model to generate 10,000 random aquifer structures. Following the method described in Section 2.3, coupled hydrogeophysical forward simulations are performed on the random aquifer structures to obtain high-fidelity concentration, head, and ERT data. Each training sample for DOCRN includes one aquifer structure, six concentration distribution fields at different time points, and one head distribution field; each training sample for HROCN contains one aquifer structure and six ERT datasets at various time points. For both datasets, 80 % of the samples are randomly selected for training, while the remaining 20 % are used for testing. Both HROCN and DOCRN use the Adam optimizer to update network parameters, with learning rate decay applied during training to prevent overfitting (decay factor set to 0.99). Both networks have the same training epochs (200) and batch size (128), with an initial learning rate of 0.001 for DOCRN and 0.0005 for HROCN. Additionally, due to the different dimensions of the data types, directly assimilating raw data makes it difficult to balance the weights in the inversion process. Therefore, normalization is applied to each type of dynamic response data:

$$\widehat{R}_m^n = \left(R_m^n - R_{min}^n\right)\left(R_{max}^n - R_{min}^n\right)^{-1} \tag{16}$$

where $\widehat{R}_m^n$ and R_m^n represent the normalized and unnormalized values of the n th type of dynamic responses on the m th grid. R_{max}^n and R_{min}^n are the

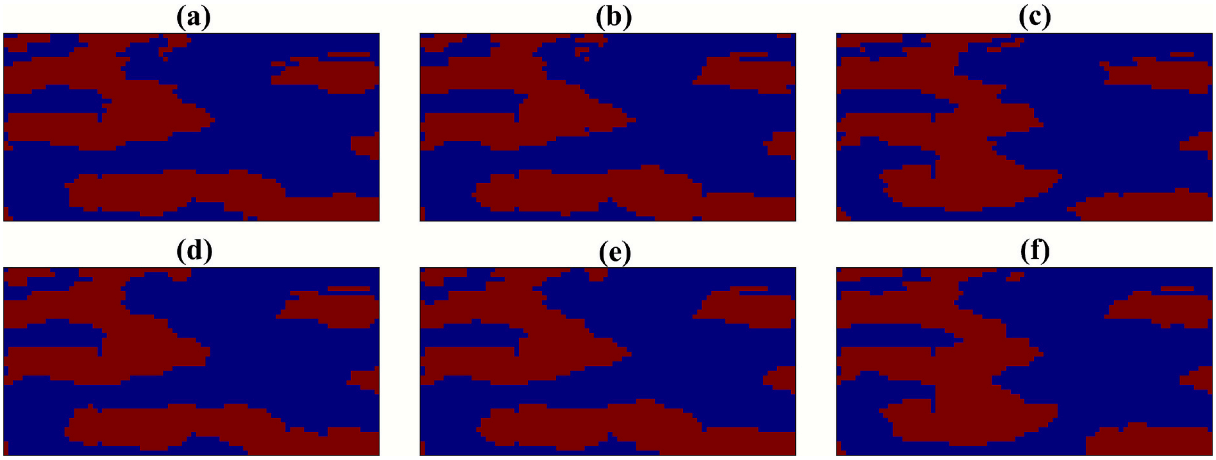


Fig. 5. Stochastic aquifer structures: (a), (b), (c); corresponding aquifer structures reconstructed by DL-based generation model: (d), (e), (f). Red represents the distribution of low-permeability facies, while blue represents the distribution of high-permeability facies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

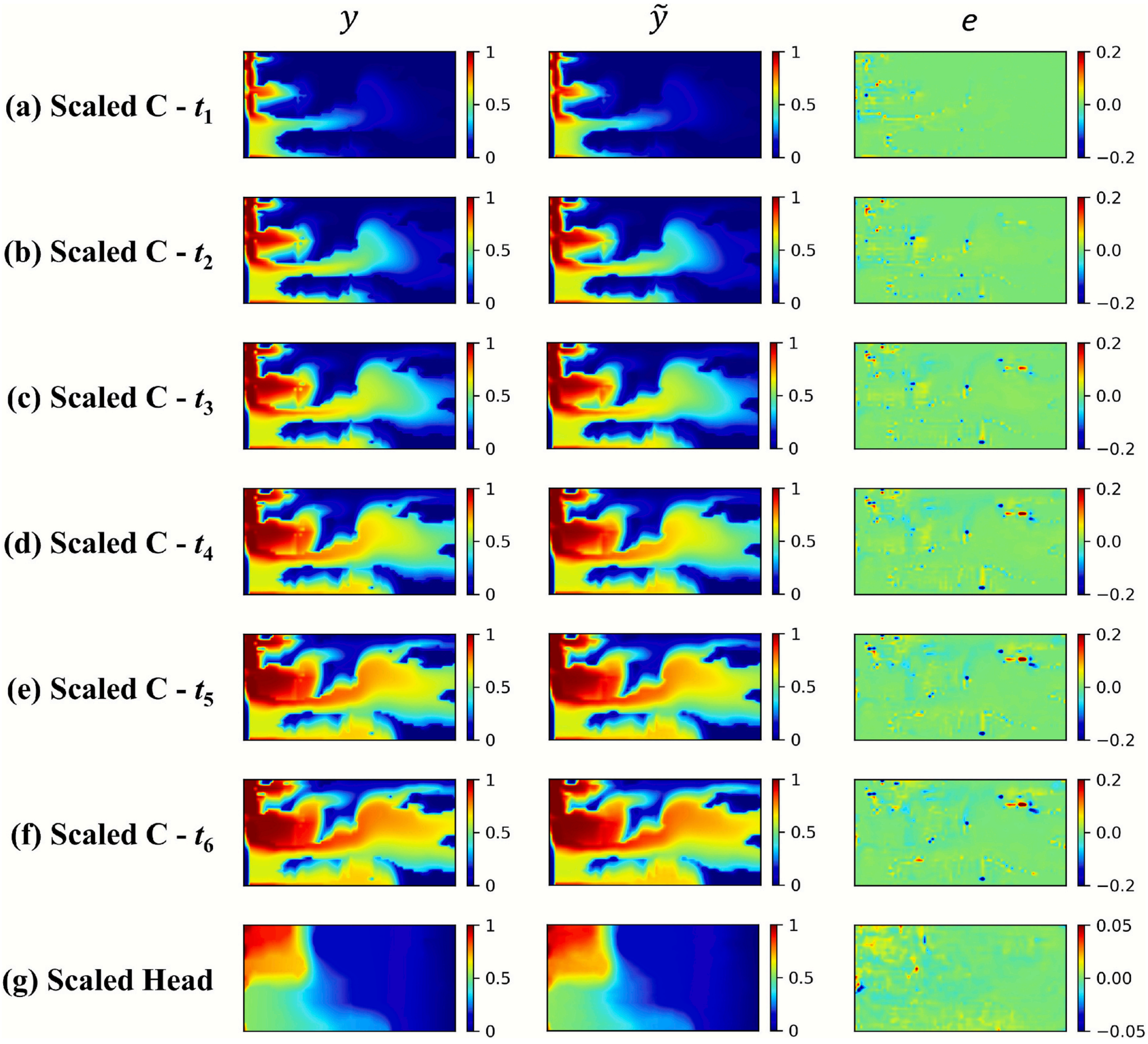


Fig. 6. Comparison of the normalized concentration fields of the contaminant plume at (a-f) six different times, Scaled C − t_1- t_6, and (g) the hydraulic head field, Scaled Head. y represents the results from the high-fidelity numerical model, $\tilde{y}$ represents the predicted results of DOCRN, and e denotes the prediction errors of between y and $\tilde{y}$.

maximum and minimum values of the n th type of dynamic response data in the training samples.

To evaluate the performance of DOCRN and HROCN, reference aquifer structures not present in the training set are input into both the high-fidelity numerical models and the surrogate models to calculate prediction errors for different types of dynamic response fields. Fig. 6a-f shows the predictive capability of DOCRN for concentration fields at various time steps, and Fig. 6g shows its performance in predicting the hydraulic head field. Note that the concentration and hydraulic head fields generated by the high-fidelity numerical model are normalized to the same scale as the DOCRN training set using Equation (16). The prediction error maps were generated through a direct comparison between DOCRN and high-fidelity numerical model outputs. The concentration field errors at all time steps are within a range of ±0.2, with larger errors mainly located at the edges of the solute plume. The error range for the hydraulic head field is within ±0.05, indicating that DOCRN's predictions for the hydraulic head field are more accurate than those for the concentration fields. In summary, DOCRN demonstrated reliable predictive capabilities in terms of accuracy compared to the high-fidelity numerical model. To obtain the dynamic response data of concentration and hydraulic head at monitoring points, time-series data at the corresponding locations are extracted from the state fields predicted by DOCRN.

4.1.3. Performance evaluation of HROCN

As described in Section 2.4, the overall network design of HROCN inherits the structural characteristics of ConvNeXt while extensively integrating the core innovation of ResNet, residual connections, to enhance the efficiency and stability of feature propagation. To systematically evaluate the performance of HROCN, this study first conducts a comparative analysis against two representative architectures, ResNet50 and ConvNeXt. All models are trained on the 10,000 sets of randomly generated structures and corresponding ERT data introduced in Section 4.1.1. The training and testing split and hyperparameter settings are kept consistent across models to ensure a fair comparison. The apparent resistivity data from ERT surveys are collected on a uniform spatial grid with contiguous measurement volumes. Therefore, this study organizes the apparent resistivity data into a 1D vector form for direct transmission within the inversion framework. The performance evaluation metrics for the three alternative models include the regularized L_1 loss function (see Supporting Information Text S2), the coefficient of determination (R^2) and the Root-Mean-Square Error ($RMSE$):

$$\begin{cases} R^2 = \dfrac{\sum_{i=1}^{n}\left(\widetilde{y}_i - \bar{y}\right)}{\sum_{i=1}^{n}(y_i - \bar{y})} \\ RMSE = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}\left(\widetilde{y}_i - y_i\right)^2} \end{cases} \tag{17}$$

where $\widetilde{y}_i$ represents the output of surrogate models, y_i is the output of the high-fidelity numerical model, $\bar{y}$ is the mean value of y_i. A lower L_1, $RMSE$ and a R^2 value approaching 1.0 suggest better performance of surrogate models. It should be emphasized that only the L_1 loss was used during training, whereas R^2 and RMSE were employed solely for post-training evaluation and comparison. The comparative results of the three models are presented in Supporting Information Text S3.

Fig. 7 shows a comparison between the predictions of HROCN and high-fidelity numerical model. The red dashed line represents the standard fit line, where the x-axis corresponds to normalized results from the high-fidelity numerical model, and the y-axis corresponds to predictions from HROCN. All scatter points are located near the standard fit line, with no significant outliers observed. The calculated R^2 is 0.9988, and the $RMSE$ is 0.0085. The comparison results indicate that HROCN can provide reliable predictions of apparent resistivity.

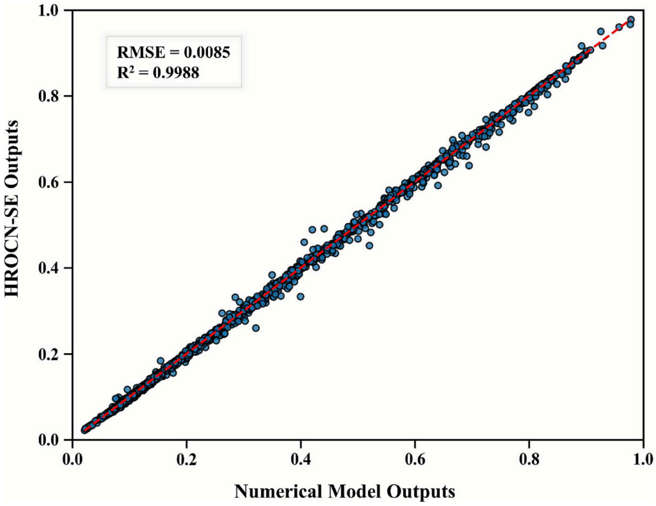


Fig. 7. Comparison of ERT data predictions between HROCN and the high-fidelity numerical model.

4.2. Inversion results

4.2.1. Accuracy of aquifer structure identification

The identification of aquifer structures is performed according to the steps described in Section 2.1. The initial sample size of latent vectors is set to 2000, and the number of ILUES iterations is 40. All cases shared the same prior parameter set, yielding 2000 posterior aquifer structures after assimilating dynamic observations for each scenario. The convergence criteria of ILUES and the sensitivity analysis of ensemble size are provided in Supporting Information Text S4. Posterior aquifer structures from different cases are compared with the facies represented by each grid in the true structure. Fig. 8 presents the distribution of aquifer structure identification accuracy and its ensemble standard deviation under different observation data scenarios. The Prior scenario, based solely on borehole data, performs poorly overall, with an average accuracy of only 81.95 %, a minimum value as low as 68.38 %, and a standard deviation of 4.41 %, indicating not only low precision but also high uncertainty. In contrast, the use of dynamic observations significantly improves identification accuracy and reduces uncertainty. Among single data types, hydraulic head (Case 1) and ERT (Case 4) both demonstrate strong performance. Case 1 achieves an average accuracy of 95.15 % with a standard deviation of only 0.12 %, indicating high stability. Case 4 attains the highest average accuracy (95.85 %) and the lowest standard deviation (0.10 %) across all scenarios, with the minimum accuracy remaining above 95 %, demonstrating that ERT provides the strongest and most stable constraint on aquifer structure. In comparison, Case 2, which uses only solute concentration data, has a lower average accuracy of 92.53 % and an increased standard deviation of 0.68 %, indicating its relatively limited independent constraint on structure identification. In joint inversion, the combination of hydraulic head and solute concentration (Case 3) results in an average accuracy of 93.09 %, higher than using concentration alone, but its standard deviation rises to 1.08 %. This indicates that while the joint data improves certain samples (with a maximum accuracy of 96.47 %, exceeding the 95.56 % achieved by hydraulic head alone), it also introduces greater variability among samples. Furthermore, when ERT data are incorporated into the joint inversion (Case 5), the average accuracy increases to 93.35 % and the standard deviation drops significantly to 0.57 %, showing that the inclusion of ERT not only enhances overall accuracy but also effectively reduces uncertainty in multi-source joint inversion.

Fig. 9 presents the average aquifer identification accuracy at different depths for different posterior scenarios, where the x-axis represents depth and the y-axis corresponds to the average accuracy across all samples in the posterior ensemble. For example, a point at 8 m on the

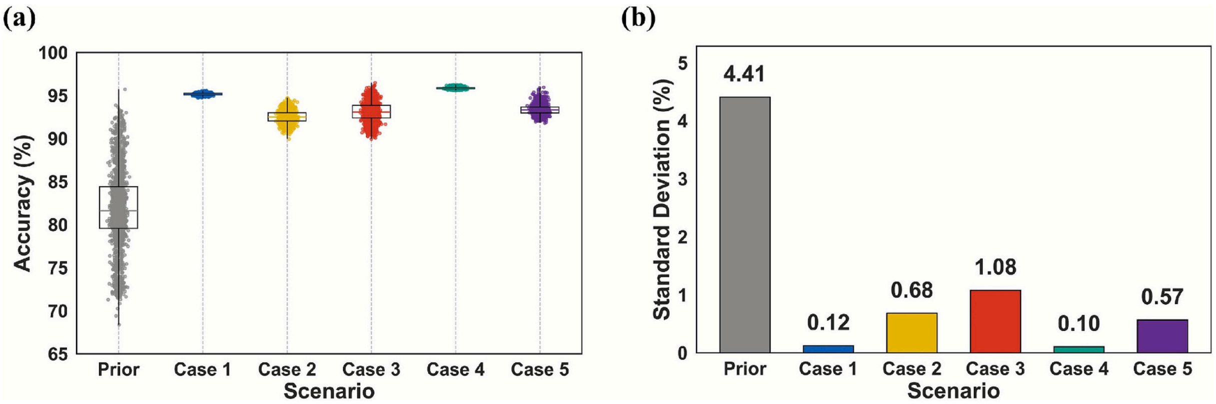


Fig. 8. (a) Boxplot of aquifer structure identification accuracy for prior and different posterior scenarios. (b) Bar chart of ensemble standard deviation for the prior and different posterior scenarios.

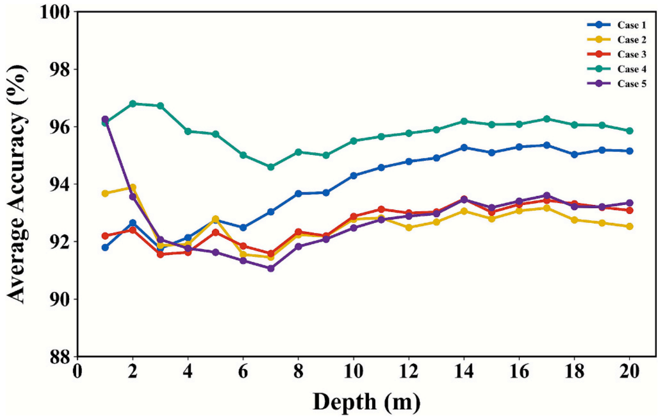


Fig. 9. Average aquifer identification accuracy at different depths for different posterior scenarios.

x-axis with a y-value of 92 % indicates that the average identification accuracy of the 2000 samples from the surface to 8 m depth is 92 %. Overall, the curves for all scenarios exhibit varying degrees of fluctuation in the shallow region (0–7 m) and gradually stabilize with a slight increase at greater depths (>10 m). This pattern is closely related to the inherent heterogeneity of the aquifer structure: in shallow regions, more complex facies variations generally reduce the accuracy of inversion results, whereas in deeper regions, the relatively continuous facies leads to more stable identification accuracy. Among the single data scenarios, ERT (Case 4) performs best, maintaining an average accuracy above 95 % across the entire profile without noticeable degradation with depth. In particular, in the shallow region where heterogeneity is most pronounced, its curve is significantly higher than those of other scenarios, indicating that the high sensitivity of ERT in shallow regions enables it to capture abrupt facies changes while still providing stable constraints at depth. In contrast, hydraulic head data (Case 1) is substantially affected in the shallow region, but its accuracy gradually increases with depth, approaching 95 % in the middle region. Concentration data (Case 2) alone performs relatively weakly, with accuracy remaining around 92 % and showing large fluctuations from shallow to the middle region. In joint inversion, the combination of hydraulic head and concentration (Case 3) achieves slightly higher accuracy at depth compared with Case 2, reflecting a certain degree of complementarity between the two data types. However, the accuracy in the shallow region remains low and fluctuations persist, indicating that multi-source coupling does not significantly mitigate the effects of heterogeneity. Notably, when ERT data are incorporated (Case 5), the average accuracy in the shallow region matches that of Case 4, but the performance in the deep region does not retain the same advantage and is comparable to Case 3. This suggests that while ERT enhances shallow region identification, its deep region advantage is attenuated in multi-source joint inversion due to the weighting and coupling interactions among different data types.

4.2.2. *Uncertainty analysis of inversion results*

The impact of different types of data on the uncertainty of aquifer structure inversion is evaluated using facies probability map and tracer breakthrough curves (BTC) from observation wells. The facies probability map in this study is obtained by calculating the sample mean of the indicator variable for the low-permeability facies:

$$p^L(k) = \frac{1}{N}\sum\nolimits_{i=1}^{N} I_i^L(k) \tag{18}$$

where $p^L(k)$ represents the probability of low-permeability facies at location k, N is the total number of samples, and $I_i^L(k)$ is the indicator variable for low-permeability facies. For the i th sample, if the facies at location k is low-permeability facies, then $I_i^L(k) = 1$; otherwise, $I_i^L(k) = 0$.

As shown in Fig. 10a, the facies probability map of the prior structural ensemble exhibits high overall uncertainty. Only at borehole (conditioning observation) locations do the occurrence probability of low-permeability facies remain stable at 0 or 1, indicating that the prior samples already incorporate conditioning constraints. After inversion with different types of observation data, the uncertainty of the structural ensemble is markedly reduced and is primarily concentrated along facies transition boundaries. A comparison of results reveals that assimilation of ERT data (Fig. 10e) produces the lowest overall uncertainty, with notably lower uncertainty in the shallow region than in the deep region, demonstrating the stronger constraint of ERT on shallow structures. In contrast, assimilation of hydraulic head data (Fig. 10b) results in lower uncertainty than assimilation of concentration data (Fig. 10c), but neither exhibits the depth-dependent pattern observed in the ERT case. This difference is related to data acquisition methods: ERT relies on surface electrodes and suffers from signal attenuation with depth, whereas hydraulic head and concentration data are collected as point measurements at fixed well locations. The joint assimilation of hydraulic head and concentration data (Fig. 10d) produces higher uncertainty than either dataset alone, suggesting that the inclusion of low-quality or weakly constraining data may compromise overall inversion performance. However, in the joint assimilation of all three data types (Fig. 10f), this adverse effect is effectively mitigated, and overall uncertainty is substantially reduced. It is noteworthy that in Figs. 10d, a clear misinterpretation occurs near the middle depth of the right boundary, where a region that should exhibit high probability (≈1) is incorrectly classified as low probability (≈0). With the incorporation of ERT data, this misinterpretation is corrected in Fig. 10f, further

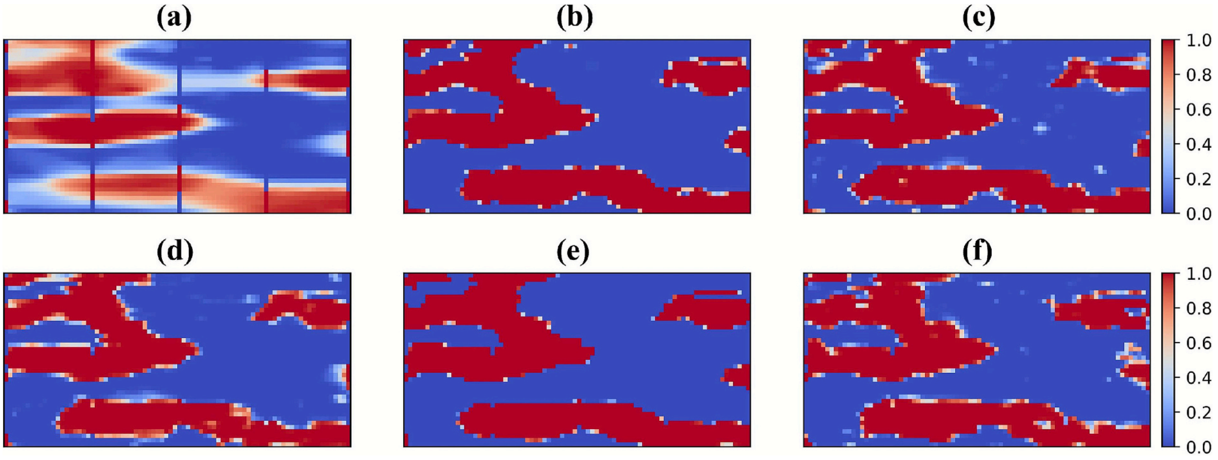


Fig. 10. The facies probability map of low-permeability facies: (a) result without inversion; (b-f) results after inversion for Case 1–5, respectively.

underscoring the critical role of ERT in reducing uncertainty and refining boundary delineation.

The BTC uncertainty analysis reflects the impact of the differences between the posterior aquifer structure and the true structure on solute transport simulation. Generally, the smaller the difference, the more accurate the solute transport simulation based on the posterior structure. The true aquifer structure, the prior aquifer structure ensemble, and the posterior aquifer structure ensembles from Case 1–5 are input into the high-fidelity numerical model to obtain dynamic responses over the simulation period. Fig. 11 shows the BTC curves at Well 3, while those at Wells 1 and 2 are provided in Figs. S5 and S6 of the Supporting Information. The orange curve represents dynamic observations from the true aquifer structure, while the blue curves represent the dynamic responses of 2000 prior and posterior structure ensembles under different scenarios. A closer match between the blue and orange curves indicates smaller heterogeneity differences between the identified and true aquifer structures. The prior ensemble results exhibit pronounced uncertainty, with wide curve dispersion and large deviations from the observed response. After inversion, assimilation of different observations improves both accuracy and uncertainty to varying degrees. Assimilating hydraulic head data (Case 1) and solute concentration data (Case 2) both effectively narrow the ensemble spread, yielding simulations closer to the observed curve. Solute concentration data inversion achieves better consistency in terms of peak concentration and decay trends, whereas hydraulic head data inversion shows a slight shift in peak arrival time. Joint assimilation of hydraulic head and solute concentration data (Case 3) alleviates this peak shift but produces a wider curve spread, suggesting higher uncertainty. Assimilation of ERT data alone (Case 4) provides strong constraints, yielding BTCs that closely match the observed response with the lowest uncertainty, indicating that ERT captures the overall dynamics of solute transport most effectively. In the joint assimilation of all three data types (Case 5), the increased uncertainty caused by hydraulic head and solute concentration data joint assimilation is partially suppressed, while the consistency in peak concentration and peak arrival time with the true response is maintained.

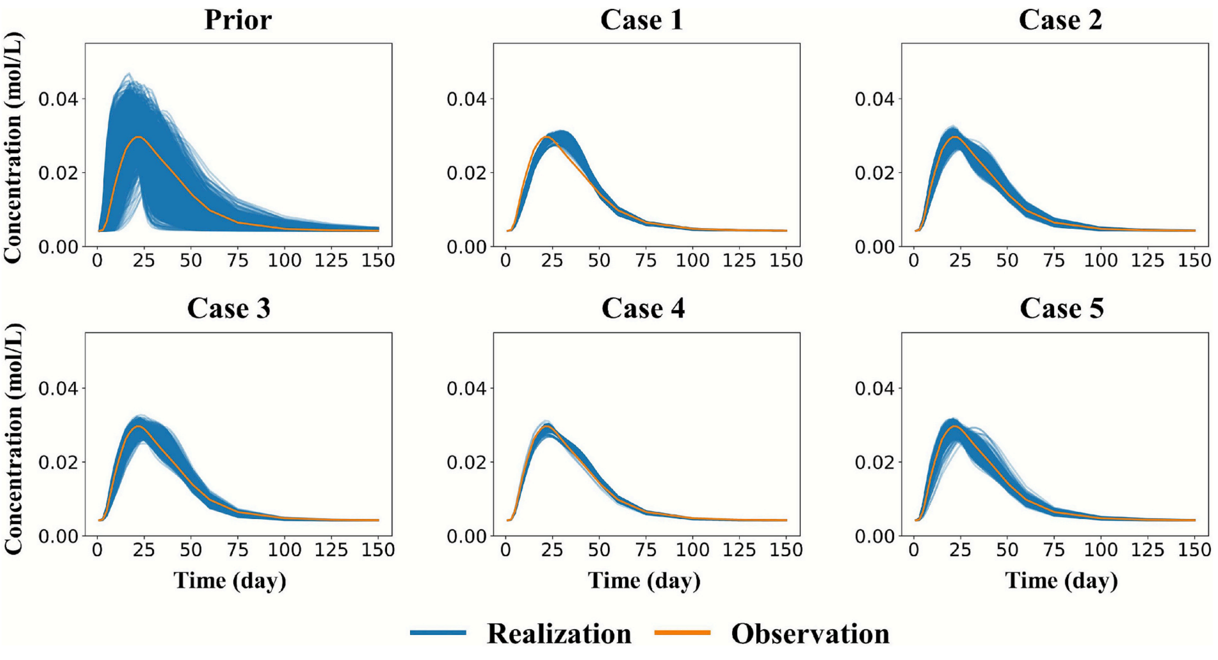


Fig. 11. Ensembles of tracer BTC curves at Well 3 for the prior and Cases 1–5. The orange curve represents the dynamic observations from the true aquifer structure, while the blue curve represents the dynamic responses of the identified aquifer structure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Discussion

4.3.1. Influence of single and multi-source data on structure identification

The data discussed in this study include facies, hydraulic head, solute concentration, and ERT observations, with inversion results varying in terms of accuracy and uncertainty depending on the data type. Facies observations, as the most direct conditioning data, directly determine the quality of the training dataset for the deep learning generation model during stochastic simulation. For example, in Fig. 10a, the uncertainty of facies distribution is significantly reduced at observation points and in their vicinity. This indicates that when facies observation points are sparse, the uncertainty of structure identification increases substantially, and facies distribution patterns may even emerge that differ considerably from the true structure. Dynamic point-based observations such as hydraulic head and solute concentration capture heterogeneity information along the flow path from the source to the observation point. However, such information carries high uncertainty, meaning that multiple structure configurations may yield responses consistent with the observations. Increasing the number of observation points can partially reduce this non-uniqueness (Mohamed et al., 2021). The inversion process generally tends to minimize the misfit between dynamic responses and observations. When the number of monitoring points is limited, the resulting structures may deviate considerably from the true heterogeneous configuration, even though the simulated dynamic responses still appear consistent with observations. As the number of monitoring points increases, the inversion is supported by more comprehensive information, which can thereby mitigate the issue of non-uniqueness to some extent. Although ERT method is discretely sampled in form, it is collected over a regular spatial grid, and each observation corresponds to a response over a continuous volume, thus presenting a semi-continuous distribution over the region. This wide spatial coverage and continuous response capability give ERT data significant advantages in identifying spatial structures, contributing to the lower uncertainty observed in ERT-based inversion results in this study.

Data fusion from multiple sources can improve the accuracy of single data inversions and mitigate the adverse effects of data scarcity in practical applications. All observations are generated based on well-defined physical processes, such as groundwater flow governed by Darcy’s law and solute transport following the advection–dispersion mechanism. Therefore, inversions using a single type of data often better recover features related to the physical process represented by that data. However, integrating results from different data types in traditional approaches often requires manual model adjustments and result corrections, which are prone to subjective bias. The HCGDIF enables joint inversion of multiple data types to obtain unified structure identification results, thus significantly reducing human intervention and improving objectivity and consistency. For example, results in this study demonstrate that the combination of hydraulic head and solute concentration data improves the accuracy of heterogeneous structure identification, while incorporating ERT data further mitigates the increased uncertainty caused by noise accumulation in multi-source data fusion. Particularly in the shallow region, the heterogeneity of the true structure is more pronounced. In this case, inversion accuracy based solely on hydraulic head and concentration data is relatively limited, whereas the integration of ERT data improves the overall reliability of structure identification.

4.3.2. Framework characteristics and applicability

Previous hydrogeophysic studies typically rely on traditional geophysical inversion techniques, where raw observations are inverted to obtain geophysical parameters and subsequently converted into spatial distributions of hydrologic parameters via petrophysical relationships (Gottschalk et al., 2017; Jiang et al., 2023; Sassen et al., 2012; Tsai et al., 2022). However, conventional inversion methods often employ isotropic regularization, resulting in overly smooth parameter fields that fail to capture strong heterogeneity. Numerous studies have shown that DL models can infer complex nonlinear relationships between raw geophysical observations and geophysical parameters. Compared to traditional Gauss-Newton inversion methods, they achieve unprecedented breakthroughs in computational accuracy and efficiency (Alvarez and Kodagoda, 2018; Jamil et al., 2024; Li et al., 2023; Tan et al., 2019). Nevertheless, these DL-based methods are typically limited in their ability to handle multi-source data fusion, as different types of data are often inverted independently and require manual calibration to reconcile the results. In contrast, the HCGDIF framework enables a process-based joint inversion of hydrologic and geophysical models. By simultaneously incorporating multiple types of observational data, HCGDIF produces unified and self-consistent results, overcoming the limitations of single-source or independent inversion approaches in capturing heterogeneity and integrating multi-source information.

HCGDIF is a flexible multi-source data fusion framework in which sub-models can be replaced according to practical requirements, enabling its application in diverse scenarios. In the Integrated Structure Generation Model developed in this study, both a transition probability-based stochastic simulation method and a deep learning approach were employed. The transition probability method essentially belongs to two-point geostatistics. Its main advantage lies in incorporating prior knowledge to enhance the geological plausibility of stochastic realizations. However, this method relies on the assumption of second-order stationarity. When borehole data are sparse or unevenly distributed, it often fails to accurately characterize spatial correlations between facies, leading to reduced simulation quality. To address this limitation, improved approaches have been proposed, such as multi-zone transition probability models (Zhu et al., 2015; Zhu et al., 2020). Moreover, when reliable training images are available and can adequately capture the true facies distribution, multiple-point statistics methods (Bai and Tahmasebi, 2020; Strebelle, 2020) can produce more accurate structural realizations. Therefore, in practical applications, the choice of stochastic simulation method should be adapted to the specific conditions.

In terms of computational efficiency, HCGDIF leverages DL-based surrogate models to accelerate the prediction of dynamic responses from facies structures. For instance, the forward modeling time for a single sample of concentration and hydraulic head is about 6 s, while for a single sample of ERT forward modeling, it is about 10 s. In the same scenario, DOCRN takes about 0.5 s, while HROCN takes about 0.3 s. Assuming the DA algorithm requires 2000 samples in each iteration and runs for 40 iterations. For the concentration and hydraulic head-based case, about 133 h are required during the forward simulation stage, while DOCRN only requires about 11 h, resulting in a 12-fold improvement in computational efficiency. For the ERT data-based case, about 222 h are required during the forward simulation stage, while HROCN only requires about 6.5 h, resulting in a 34-fold improvement in computational efficiency.

Furthermore, HCGDIF demonstrates strong extensibility. In more complex geological settings, such as aquifer systems with multiple facies zones or multi-scale heterogeneity, the transition probability model can be expanded by adjusting structural parameters to represent different facies patterns (Soltanian et al., 2015). After experimentally measuring the mean and variance of permeability coefficients for different facies types, the aquifer structure model can be transformed into a heterogeneous permeability field using the Karhunen-Loève expansion method. The HROCN proposed in this study also demonstrates strong generalization capability and can be extended to other geophysical methods such as self-potential (Kang et al., 2020; Soueid Ahmed et al., 2016) and induced polarization (Römhild et al., 2024), to improve data prediction accuracy and computational efficiency.

4.3.3. Research limitations

Despite the promising potential of HCGDIF, the use of synthetic experiments in this study limits its generalizability, and several challenges remain in practical applications. (1) The uncertainty of petrophysical relationships remains a key factor affecting the accuracy of coupled hydrogeophysic modeling (Brunetti and Linde, 2018; Friedli and Linde, 2024). In this study, Archie's law was employed, which is applicable to clay-free heterogeneous media. For clay-rich media, the effect of cations sorbed on the mineral surfaces must be additionally considered (Kang et al., 2019; Revil et al., 2018). Some studies have also constructed empirical relationship models through laboratory experiments (Binley et al., 2015; Linde et al., 2017). However, in complex field environments, the applicability and stability of these relationships and their parameters are highly uncertain. (2) Observation data inevitably contain noise from various sources, including instrument limitations, operational errors, and environmental interference (Huisman et al., 2010; Linde et al., 2015). To make the synthetic experiments more realistic, a relatively high level of noise was introduced into the simulated ERT data. Nevertheless, in single data inversion scenarios, ERT data still yielded the best results, indicating a certain robustness to observational noise. It is noteworthy that, in theory, the reliability of inversion results should increase with the amount of observational data. However, this study found that the inversion using the limited hydraulic head data outperformed that using solute concentration data. This can be attributed to the fact that while DL-based surrogate models substantially enhance computational efficiency, they inevitably introduce model noise. Such noise can be interpreted as the deviation between surrogate model predictions and high-fidelity numerical simulations. On the one hand, the generalization ability of deep neural networks depends on the quality and diversity of training data. The model may produce prediction bias when data is insufficient or feature extraction is inadequate; on the other hand, the design of the network architecture and optimization strategies may also introduce new uncertainties, such as overfitting, gradient vanishing or explosion, etc., which further affect prediction accuracy (Chan and Elsheikh, 2019; Mo et al., 2020; Siade et al., 2020; Xiao et al., 2021). Therefore, a systematic investigation into the mechanisms of observation and model noise under different data types, along with the development of compensation and correction strategies, represents a crucial direction for extending HCGDIF to practical applications.

5. Conclusions

This study develops a Combined Hydrologic, Chemical, and Geophysical Data Deep Inversion Framework (HCGDIF) for the identification of heterogeneous aquifer structures under varying data availability conditions. The framework consists of three core components: (1) the Integrated Structure Generation Model for dimensionality reduction of heterogeneous structures; (2) deep learning (DL)-based surrogate models for predicting hydrogeophysical dynamic responses; and (3) the ILUES algorithm for minimizing the mismatch between observations and simulated responses. Based on a synthetic electrical resistivity tomography (ERT)-monitored tracer experiment, the following key findings are obtained:

1. The Integrated structure generation model can generate a large number of random structures from limited facies observations and parameterize heterogeneous structures into latent vectors of customizable dimensions. This effectively addresses the challenge of inverting high-dimensional heterogeneous parameters and provides sufficient training data for the DL models.
2. The DL-based surrogate models achieve rapid prediction of dynamic responses while maintaining accuracy comparable to high-fidelity numerical simulations. In particular, the Hybrid Residual Octave-Convolution network with Squeeze-and-Excitation attention (HROCN), which integrates octave convolution layers, dense residual connections, depthwise separable convolutions, and attention mechanisms, effectively captures the complex nonlinear mapping between aquifer structures and ERT data, yielding higher accuracy than baseline models.
3. Due to differences in sampling strategies and underlying physical processes, different data types contribute differently to structure identification. When hydraulic head and solute concentration data are sparse, their identification accuracy is lower than that of ERT data. Although multi-source data fusion provides higher identification accuracy than single data, it may also lead to greater uncertainty due to the accumulation of observation noise. The inclusion of ERT data, with its regionally semi-continuous distribution characteristics, helps reduce this uncertainty.

The scenarios considered in this study include independent and joint inversions using hydraulic head, solute concentration, and ERT data. All results clearly demonstrate the robustness of the HCGDIF under various data conditions. In addition, the HCGDIF achieves synergistic inversion of hydrologic and geophysical models and significantly improves inversion efficiency through the incorporation of DL techniques. In future practical applications, beyond the differences in the information content carried by various types of data, data errors and missing observations may also affect inversion performance to varying degrees. The findings of this study highlight the importance of integrating multiple complementary observation datasets, particularly those with broader spatial coverage, to enhance the robustness of structure identification. Future research should place greater emphasis on optimizing monitoring networks, evaluating noise propagation across multi-source datasets, and improving strategies for fusing data in real-world applications.

CRediT authorship contribution statement

Yuzhou Xia: Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Chuanjun Zhan:** Writing – review & editing, Visualization, Methodology, Conceptualization. **Zhenxue Dai:** Writing – review & editing, Project administration, Funding acquisition, Conceptualization. **Jichun Wu:** Writing – review & editing. **Xiaoying Zhang:** Writing – review & editing. **Huichao Yin:** Writing – review & editing. **Lin Zhu:** Writing – review & editing. **Jiahe Yan:** Writing – review & editing, Methodology. **Zihao Wang:** Writing – review & editing. **Mohamad Reza Soltanian:** Writing – review & editing. **Kenneth C. Carroll:** Writing – review & editing.

Declaration of competing interest

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

Acknowledgments

This work is funded by the National Natural Science Foundation of China (NSFC: U2267217, 42141011, 42002254), Shandong Key Water Conservancy Science and Technology Project (2024370203001957) and Shandong Provincial Natural Science Foundation (ZR2024QD095).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2025.134701.

Data availability

Data will be made available on request.

References

Alvarez, J.K., Kodagoda, S., 2018. Application of deep learning image-to-image transformation networks to GPR radargrams for sub-surface imaging in infrastructure monitoring. In: 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 611–616. https://doi.org/10.1109/ICIEA.2018.8397788.

Bai, T., Tahmasebi, P., 2020. Hybrid geological modeling: Combining machine learning and multiple-point statistics. Comput. Geosci. 142. https://doi.org/10.1016/j.cageo.2020.104519.

Bear, J., 1979. Hydraulics of Groundwater. McGraw-Hill International Book Company, p. 569.

Berg, S.J., Illman, W.A., 2011. Capturing aquifer heterogeneity: Comparison of approaches through controlled sandbox experiments. Water Resour. Res. 47 (9). https://doi.org/10.1029/2011WR010429.

Bianchi, M., Pedretti, D., 2018. An entrogram-based approach to describe spatial heterogeneity with applications to solute transport in porous media. Water Resour. Res. 54 (7), 4432–4448. https://doi.org/10.1029/2018wr022827.

Binley, A., et al., 2015. The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. Water Resour. Res. 51 (6), 3837–3866. https://doi.org/10.1002/2015wr017016.

Bravo, H.R., Jiang, F., Hunt, R.J., 2002. Using groundwater temperature data to constrain parameter estimation in a groundwater flow model of a wetland system. Water Resour. Res. 38 (8). https://doi.org/10.1029/2000wr000172.

Brunetti, C., Bianchi, M., Pirot, G., Linde, N., 2019. Hydrogeological model selection among complex spatial priors. Water Resour. Res. 55 (8), 6729–6753. https://doi.org/10.1029/2019wr024840.

Brunetti, C., Linde, N., 2018. Impact of petrophysical uncertainty on Bayesian hydrogeophysical inversion and model selection. Adv. Water Resour. 111, 346–359. https://doi.org/10.1016/j.advwatres.2017.11.028.

Camporese, M., Cassiani, G., Deiana, R., Salandin, P., Binley, A., 2015. Coupled and uncoupled hydrogeophysical inversions using ensemble Kalman filter assimilation of ERT-monitored tracer test data. Water Resour. Res. 51 (5), 3277–3291. https://doi.org/10.1002/2014wr016017.

Canchumuni, S.A., Emerick, A.A., Pacheco, M.A., 2017. Integration of ensemble data assimilation and deep learning for history matching facies models. OTC Brasil. https://doi.org/10.4043/28015-ms.

Cardiff, M., Barrash, W., Kitanidis, P.K., 2013. Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. Water Resour. Res. 49 (11), 7311–7326. https://doi.org/10.1002/wrcr.20519.

Carle, S.F., Fogg, G.E., 1996. Transition probability-based indicator geostatistics. Math. Geol. 28 (4), 453–476. https://doi.org/10.1007/BF02083656.

Chan, S., Elsheikh, A.H., 2019. Parametric generation of conditional geological realizations using generative neural networks. Comput. Geosci. 23 (5), 925–952. https://doi.org/10.1007/s10596-019-09850-7.

Chen, C., et al., 2024. Deep learning-based inversion framework by assimilating hydrogeological and geophysical data for an enhanced geothermal system characterization and thermal performance prediction. Energy 302. https://doi.org/10.1016/j.energy.2024.131713.

Chen, J., et al., 2022. Integration of deep learning and information theory for designing monitoring networks in heterogeneous aquifer systems. Water Resour. Res. 58 (10). https://doi.org/10.1029/2022wr032429.

Chen, J., Dai, Z., Yin, S., Zhang, M., Soltanian, M.R., 2025. Enhancing inverse modeling in groundwater systems through machine learning: a comprehensive comparative study. Hydrol. Earth Syst. Sci. 29 (17), 4251–4279. https://doi.org/10.5194/hess-29-4251-2025.

Chen, J., Rubin, Y., 2003. An effective Bayesian model for lithofacies estimation using geophysical data. Water Resour. Res. 39 (5). https://doi.org/10.1029/2002wr001666.

Chen, Y., et al., 2019. Drop an octave: reducing spatial redundancy in convolutional neural networks with octave convolution. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 3434–3443. https://doi.org/10.1109/iccv.2019.00353.

Christensen, N.K., Christensen, S., Ferre, T.P.A., 2016. Testing alternative uses of electromagnetic data to reduce the prediction error of groundwater models. Hydrol. Earth Syst. Sci. 20 (5), 1925–1946. https://doi.org/10.5194/hess-20-1925-2016.

Cirpka, O.A., Stettler, M.M., Dentz, M., 2022. Spatial Markov model for the prediction of travel-time-based solute dispersion in three-dimensional heterogeneous media. Water Resour. Res. 58 (6). https://doi.org/10.1029/2022wr032215.

Cockett, R., Kang, S., Heagy, L.J., Pidlisecky, A., Oldenburg, D.W., 2015. SimPEG: an open source framework for simulation and gradient based parameter estimation in geophysical applications. Comput. Geosci. 85, 142–154. https://doi.org/10.1016/j.cageo.2015.09.015.

Dai, Z., et al., 2014. An integrated framework for optimizing CO2 sequestration and enhanced oil recovery. Environ. Sci. Technol. Lett. 1 (1), 49–54. https://doi.org/10.1021/ez4001033.

Dai, Z., Wolfsberg, A., Lu, Z., Ritzi, R., 2007. Representing aquifer architecture in macrodispersivity models with an analytical solution of the transition probability matrix. Geophys. Res. Lett. 34 (20). https://doi.org/10.1029/2007gl031608.

Dai, Z., et al., 2020. How does resolution of sedimentary architecture data affect plume dispersion in multiscale and hierarchical systems? J. Hydrol. 582. https://doi.org/10.1016/j.jhydrol.2019.124516.

Dausman, A.M., Doherty, J., Langevin, C.D., Sukop, M.C., 2010. Quantifying data worth toward reducing predictive uncertainty. Groundwater 48 (5), 729–740. https://doi.org/10.1111/j.1745-6584.2010.00679.x.

Deng, J. et al., 2009. Imagenet: A large-scale hierarchical image database, 2009 IEEE conference on computer vision and pattern recognition. Ieee, pp. 248-255, 10.1109/CVPR.2009.5206848.

Deng, Y., et al., 2024. Characterization of discrete fracture networks with deep-learning based hydrogeophysical inversion. J. Hydrol. 631. https://doi.org/10.1016/j.jhydrol.2024.130819.

Friedli, L., Linde, N., 2024. Solving geophysical inversion problems with intractable likelihoods: linearized gaussian approximations versus the correlated pseudo-marginal method. Math. Geosci. 56 (1), 55–75. https://doi.org/10.1007/s11004-023-10064-y.

Gershenzon, N.I., et al., 2015. Influence of small-scale fluvial architecture on CO2 trapping processes in deep brine reservoirs. Water Resour. Res. 51 (10), 8240–8256. https://doi.org/10.1002/2015wr017638.

Gottschalk, I.P., et al., 2017. Integrating non-colocated well and geophysical data to capture subsurface heterogeneity at an aquifer recharge and recovery site. J. Hydrol. 555, 407–419. https://doi.org/10.1016/j.jhydrol.2017.10.028.

Han, Z., Kang, X., Wu, J., Shi, X., 2022. Characterization of the non-Gaussian hydraulic conductivity field via deep learning-based inversion of hydraulic-head and self-potential data. J. Hydrol. 610. https://doi.org/10.1016/j.jhydrol.2022.127830.

Han, Z., Kang, X., Wu, J., Shi, X., Jiang, J., 2024. Improved solute transport modeling through joint estimation of hydraulic conductivity and dispersivities from tracer and ERT data. Adv. Water Resour. 185. https://doi.org/10.1016/j.advwatres.2024.104655.

Hart, B.S., 1999. Definition of subsurface stratigraphy, structure and rock properties from 3-D seismic data. Earth Sci. Rev. 47 (3), 189–218. https://doi.org/10.1016/S0012-8252(99)00029-X.

Heagy, L.J., Cockett, R., Kang, S., Rosenkjaer, G.K., Oldenburg, D.W., 2017. A framework for simulation and inversion in electromagnetics. Comput. Geosci. 107, 1–19. https://doi.org/10.1016/j.cageo.2017.06.018.

Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-excitation networks, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132-7141, 10.1109/CVPR.2018.00745.

Hubbard, S.S., et al., 2001. Hydrogeological characterization of the south oyster bacterial transport site using geophysical data. Water Resour. Res. 37 (10), 2431–2456. https://doi.org/10.1029/2001wr000279.

Huisman, J.A., Rings, J., Vrugt;, J.A., Sorg, J., Vereecken, H., 2010. Hydraulic properties of a model dike from coupled Bayesian and multi-criteria hydrogeophysical inversion. J. Hydrol. 380 (1), 62–73. https://doi.org/10.1016/j.jhydrol.2009.10.023.

Ikard, S.J., Carroll, K.C., Rucker, D.F., Adams, R.F., Brooks, S.C., 2023. Geoelectric characterization of hyporheic exchange flow in the bedrock-lined streambed of east fork poplar creek, oak ridge, Tennessee. Geophys. Res. Lett. 50 (8). https://doi.org/10.1029/2022gl102616.

Jamil, A., et al., 2024. Comparison of machine learning and electrical resistivity arrays to inverse modeling for locating and characterizing subsurface targets. J. Appl. Geophys. 229. https://doi.org/10.1016/j.jappgeo.2024.105493.

Jardani, A., Revil, A., Dupont, J.P., 2013. Stochastic joint inversion of hydrogeophysical data for salt tracer test monitoring and hydraulic conductivity imaging. Adv. Water Resour. 52, 62–77. https://doi.org/10.1016/j.advwatres.2012.08.005.

Jia, S., et al., 2023. Upscaling dispersivity for conservative solute transport in naturally fractured media. Water Res. 235. https://doi.org/10.1016/j.watres.2023.119844.

Jiang, Z., Ringel, L.M., Bayer, P., Xu, T., 2023. Fracture network characterization in reservoirs by joint inversion of microseismicity and thermal breakthrough data: method development and verification. Water Resour. Res. 59 (9). https://doi.org/10.1029/2022wr034339.

Johnson, T., et al., 2015. Four-dimensional electrical conductivity monitoring of stage-driven river water intrusion: accounting for water table effects using a transient mesh boundary and conditional inversion constraints. Water Resour. Res. 51 (8), 6177–6196. https://doi.org/10.1002/2014wr016129.

Jose, S.C., Rahman, M.A., Cirpka, O.A., 2004. Large-scale sandbox experiment on longitudinal effective dispersion in heterogeneous porous media. Water Resour. Res. 40 (12). https://doi.org/10.1029/2004wr003363.

Kang, X., et al., 2021. Hydrogeophysical characterization of nonstationary DNAPL source zones by integrating a convolutional variational autoencoder and ensemble smoother. Water Resour. Res. 57 (2). https://doi.org/10.1029/2020wr028538.

Kang, X., et al., 2020. Improved characterization of DNAPL source zones via sequential hydrogeophysical inversion of hydraulic-head, self-potential and partitioning tracer data. Water Resour. Res. 56 (8). https://doi.org/10.1029/2020wr027627.

Kang, X., et al., 2018. Coupled hydrogeophysical inversion of DNAPL source zone architecture and permeability field in a 3D heterogeneous sandbox by assimilation time-lapse cross-borehole electrical resistivity data via ensemble Kalman filtering. J. Hydrol. 567, 149–164. https://doi.org/10.1016/j.jhydrol.2018.10.019.

Kang, X., et al., 2019. Coupled hydrogeophysical inversion to identify non-Gaussian hydraulic conductivity field by jointly assimilating geochemical and time-lapse geophysical data. J. Hydrol. 578. https://doi.org/10.1016/j.jhydrol.2019.124092.

Karras, T., Laine, S., Aila, T., 2021. A style-based generator architecture for generative adversarial networks. IEEE Trans. Pattern Anal. Mach. Intell. 43 (12), 4217–4228. https://doi.org/10.1109/TPAMI.2020.2970919.

Khambhammettu, P., Renard, P., Doherty, J., 2020. The traveling pilot point method. A novel approach to parameterize the inverse problem for categorical fields. Adv. Water Resour. 138. https://doi.org/10.1016/j.advwatres.2020.103556.

Kim, S., Min, B., Lee, K., Jeong, H., 2018. Integration of an iterative update of sparse geologic dictionaries with ES-MDA for history matching of channelized reservoirs. Geofluids 2018, 1–21. https://doi.org/10.1155/2018/1532868.

LaBrecque, D.J., Miletto, M., Daily, W., Ramirez, A., Owen, E., 1996. The effects of noise on Occam's inversion of resistivity tomography data. Geophysics 61 (2), 538–548. https://doi.org/10.1190/1.1443980.

Laloy, E., Hérault, R., Jacques, D., Linde, N., 2018. Training-image based geostatistical inversion using a spatial generative adversarial neural network. Water Resour. Res. 54 (1), 381–406. https://doi.org/10.1002/2017wr022148.

Laloy, E., Hérault, R., Lee, J., Jacques, D., Linde, N., 2017. Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network. Adv. Water Resour. 110, 387–405. https://doi.org/10.1016/j.advwatres.2017.09.029.

Li, Z., et al., 2023. Study on rapid inversion of soil water content from ground-penetrating radar data based on deep learning. Remote Sens. (Basel) 15 (7). https://doi.org/10.3390/rs15071906.

Lin, T.-Y. et al., 2014. Microsoft coco: Common objects in context, Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13. Springer, pp. 740-755, 10.48550/arXiv.1405.0312.

Linde, N., Ginsbourger, D., Irving, J., Nobile, F., Doucet, A., 2017. On uncertainty quantification in hydrogeology and hydrogeophysics. Adv. Water Resour. 110, 166–181. https://doi.org/10.1016/j.advwatres.2017.10.014.

Linde, N., Renard, P., Mukerji, T., Caers, J., 2015. Geological realism in hydrogeological and geophysical inverse modeling: a review. Adv. Water Resour. 86, 86–101. https://doi.org/10.1016/j.advwatres.2015.09.019.

Liu, Y., Sun, W., Durlofsky, L.J., 2019. A deep-learning-based geological parameterization for history matching complex models. Math. Geosci. 51 (6), 725–766. https://doi.org/10.1007/s11004-019-09794-9.

Liu, Z. et al., 2021. Swin transformer: Hierarchical vision transformer using shifted windows, Proceedings of the IEEE/CVF international conference on computer vision, pp. 10012-10022, 10.1109/ICCV48922.2021.00986.

Liu, Z. et al., 2022. A convnet for the 2020s, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 11976-11986, 10.1109/CVPR52688.2022.01167.

Lopez-Alvis, J., Nguyen, F., Looms, M.C., Hermans, T., 2022. Geophysical inversion using a variational autoencoder to model an assembled spatial prior uncertainty. J. Geophys. Res. Solid Earth 127 (3). https://doi.org/10.1029/2021jb022581.

Ma, F., et al., 2024. Incorporating cross-scale insights into colloid-facilitated radionuclide transport in fractured rocks: a critical review. Earth Sci. Rev. 259. https://doi.org/10.1016/j.earscirev.2024.104974.

Magalhães, A.J.C., et al., 2020. High-resolution sequence stratigraphy applied to reservoir zonation and characterisation, and its impact on production performance - shallow marine, fluvial downstream, and lacustrine carbonate settings. Earth Sci. Rev. 210. https://doi.org/10.1016/j.earscirev.2020.103325.

Mao, D., et al., 2022. Infiltration assessments on top of Yungang grottoes by time-lapse electrical resistivity tomography. Hydrology 9 (5). https://doi.org/10.3390/hydrology9050077.

Mao, D., et al., 2024. Heterogeneous reservoir seepage characterized by geophysical, hydrochemical, and hydrologic methods. Geophysics 89 (5), B401–B413. https://doi.org/10.1190/geo2024-0199.1.

McGarr, J.T., Wallace, C.D., Ntarlagiannis, D., Sturmer, D.M., Soltanian, M.R., 2021. Geophysical mapping of hyporheic processes controlled by sedimentary architecture within compound bar deposits. Hydrol. Process. 35 (9). https://doi.org/10.1002/hyp.14358.

McKay, M.D., Beckman, R.J., Conover, W.J., 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 42 (1), 55–61. https://doi.org/10.2307/1268522.

Mo, S., Zabaras, N., Shi, X., Wu, J., 2020. Integration of adversarial autoencoders with residual dense convolutional networks for estimation of non-gaussian hydraulic conductivities. Water Resour. Res. 56 (2). https://doi.org/10.1029/2019wr026082.

Mohamed, R.A.M., et al., 2021. Geostatistical interpolation of streambed hydrologic attributes with addition of left censored data and anisotropy. J. Hydrol. 599. https://doi.org/10.1016/j.jhydrol.2021.126474.

Neven, A., Renard, P., 2023. A novel methodology for the stochastic integration of geophysical and hydrogeological data in geologically consistent models. Water Resour. Res. 59 (7). https://doi.org/10.1029/2023wr034992.

Pollock, D., Cirpka, O.A., 2010. Fully coupled hydrogeophysical inversion of synthetic salt tracer experiments. Water Resour. Res. 46 (7). https://doi.org/10.1029/2009wr008575.

Pollock, D., Cirpka, O.A., 2012. Fully coupled hydrogeophysical inversion of a laboratory salt tracer experiment monitored by electrical resistivity tomography. Water Resour. Res. 48 (1). https://doi.org/10.1029/2011wr010779.

Rakotonirina, N.C., Rasoanaivo, A., 2020. ESRGAN+ : further improving enhanced super-resolution generative adversarial network. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3637–3641. https://doi.org/10.1109/ICASSP40776.2020.9054071.

Reuschen, S., Jobst, F., Nowak, W., 2021. Efficient discretization-independent bayesian inversion of high-dimensional multi-gaussian priors using a hybrid MCMC. Water Resour. Res. 57 (8). https://doi.org/10.1029/2021wr030051.

Revil, A., et al., 2018. Electrical conductivity and induced polarization investigations at Krafla volcano, Iceland. J. Volcanol. Geoth. Res. 368, 73–90. https://doi.org/10.1016/j.jvolgeores.2018.11.008.

Römhild, L., Fiandaca, G., Bayer, P., 2024. Joint inversion of induced polarization and hydraulic tomography data for hydraulic conductivity imaging. Geophys. J. Int. 238 (2), 960–973. https://doi.org/10.1093/gji/ggae197.

Rucker, D.F., et al., 2021. Bedrock architecture, soil texture, and hyporheic zone characterization combining electrical resistivity and induced polarization imaging. J. Appl. Geophys. 188. https://doi.org/10.1016/j.jappgeo.2021.104306.

Sassen, D.S., et al., 2012. Reactive facies: an approach for parameterizing field-scale reactive transport models using geophysical methods. Water Resour. Res. 48 (10). https://doi.org/10.1029/2011wr011047.

Sen, P.N., Goode, P.A., 1992. Influence of temperature on electrical conductivity on shaly sands. Geophysics 57 (1), 89–96. https://doi.org/10.1190/1.1443191.

Siade, A.J., Cui, T., Karelse, R.N., Hampton, C., 2020. Reduced-dimensional gaussian process machine learning for groundwater allocation planning using swarm theory. Water Resour. Res. 56 (3). https://doi.org/10.1029/2019wr026061.

Soltanian, M.R., Ritzi, R.W., 2014. A new method for analysis of variance of the hydraulic and reactive attributes of aquifers as linked to hierarchical and multiscaled sedimentary architecture. Water Resour. Res. 50 (12), 9766–9776. https://doi.org/10.1002/2014wr015468.

Soltanian, M.R., Ritzi, R.W., Huang, C.C., Dai, Z., 2015. Relating reactive solute transport to hierarchical and multiscale sedimentary architecture in a Lagrangian-based transport model: 2. Particle Displacement Variance. Water Resour. Res. 51 (3), 1601–1618. https://doi.org/10.1002/2014wr016354.

Song, X., et al., 2019. Delineating facies spatial distribution by integrating ensemble data assimilationand indicator geostatistics with level-set transformation. Water Resour. Res. 55 (4), 2652–2671. https://doi.org/10.1029/2018wr023262.

Soueid Ahmed, A., Jardani, A., Revil, A., Dupont, J.P., 2014. Hydraulic conductivity field characterization from the joint inversion of hydraulic heads and self-potential data. Water Resour. Res. 50 (4), 3502–3522. https://doi.org/10.1002/2013wr014645.

Soueid Ahmed, A., Jardani, A., Revil, A., Dupont, J.P., 2016. Joint inversion of hydraulic head and self-potential data associated with harmonic pumping tests. Water Resour. Res. 52 (9), 6769–6791. https://doi.org/10.1002/2016wr019058.

Strebelle, S., 2020. Multiple-point statistics simulation models: pretty pictures or decision-making tools? Math. Geosci. 53 (2), 267–278. https://doi.org/10.1007/s11004-020-09908-8.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning, Proceedings of the AAAI conference on artificial intelligence, 10.48550/arXiv.1602.07261.

Tan, C., Lv, S., Dong, F., Takei, M., 2019. Image reconstruction based on convolutional neural network for electrical resistance tomography. IEEE Sens. J. 19 (1), 196–204. https://doi.org/10.1109/jsen.2018.2876411.

Tsai, C.H., Rucker, D.F., Brooks, S.C., Ginn, T., Carroll, K.C., 2022. Transient storage model parameter optimization using the simulated annealing method. Water Resour. Res. 58 (7), e2022WR032018. https://doi.org/10.1029/2022WR032018.

Wallace, C.D., Soltanian, M.R., 2021. Underlying riparian lithology controls redox dynamics during stage-driven mixing. J. Hydrol. 595. https://doi.org/10.1016/j.jhydrol.2021.126035.

Wang, N., Chang, H., Zhang, D., 2021. Deep-learning-based inverse modeling approaches: a subsurface flow example. J. Geophys. Res. Solid Earth 126 (2). https://doi.org/10.1029/2020jb020549.

Xia, X., Jiang, S., Zhou, N., Cui, J., Li, X., 2023. Groundwater contamination source identification and high-dimensional parameter inversion using residual dense convolutional neural network. J. Hydrol. 617. https://doi.org/10.1016/j.jhydrol.2022.129013.

Xiang, J., Yeh, T.-C.J., Lee, C.-H., Hsu, K.-C., Wen, J.-C., 2009. A simultaneous successive linear estimator and a guide for hydraulic tomography analysis. Water Resour. Res. 45 (2). https://doi.org/10.1029/2008WR007180.

Xiao, C., Leeuwenburgh, O., Lin, H.-X., Heemink, A., 2021. Conditioning of deep-learning surrogate models to image data with application to reservoir characterization. Knowl.-Based Syst. 220. https://doi.org/10.1016/j.knosys.2021.106956.

Xu, L., et al., 2025. Fractal and NMR based characterizations of multi-scale pore structure alterations in tight sandstones due to scCO2-water–rock interactions. Chem. Eng. J. 508. https://doi.org/10.1016/j.cej.2025.160898.

Xu, T., Gómez-Hernández, J.J., Chen, Z., Lu, C., 2021. A comparison between ES-MDA and restart EnKF for the purpose of the simultaneous identification of a contaminant source and hydraulic conductivity. J. Hydrol. 595. https://doi.org/10.1016/j.jhydrol.2020.125681.

Xu, T., Gómez-Hernández, J.J., 2016. Characterization of non-Gaussian conductivities and porosities with hydraulic heads, solute concentrations, and water temperatures. Water Resour. Res. 52 (8), 6111–6136. https://doi.org/10.1002/2016wr019011.

Xu, T., Sonnenthal, E., Spycher, N., Pruess, K., 2006. TOUGHREACT—A simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: applications to geothermal injectivity and CO2 geological sequestration. Comput. Geosci. 32 (2), 145–165. https://doi.org/10.1016/j.cageo.2005.06.014.

Yeh, T.-C.J., Lee, C.-H., Hsu, K.-C., Wen, J.-C., 2008. Fusion of hydrologic and geophysical tomographic surveys. Geosci. J. 12 (2), 159–167. https://doi.org/10.1007/s12303-008-0017-6.

Yeh, T.-C.J., Liu, S., 2000. Hydraulic tomography: development of a new aquifer test method. Water Resour. Res. 36 (8), 2095–2105. https://doi.org/10.1029/2000WR900114.

Yeh, T.-C.J., Šimůnek, J., 2002. Stochastic fusion of information for characterizing and monitoring the vadose zone. Vadose Zone J. 1 (2), 207–221. https://doi.org/10.2136/vzj2002.2070.

Zhan, C., et al., 2022. An integrated inversion framework for heterogeneous aquifer structure identification with single-sample generative adversarial network. J. Hydrol. 610. https://doi.org/10.1016/j.jhydrol.2022.127844.

Zhan, C., Dai, Z., Soltanian, M.R., de Barros, F.P.J., 2022. Data-worth analysis for heterogeneous subsurface structure identification with a stochastic deep learning framework. Water Resour. Res. 58 (11). https://doi.org/10.1029/2022wr033241.

Zhan, C., Dai, Z., Soltanian, M.R., Zhang, X., 2021. Stage-wise stochastic deep learning inversion framework for subsurface sedimentary structure identification. Geophys. Res. Lett. 49 (1). https://doi.org/10.1029/2021gl095823.

Zhang, J., Lin, G., Li, W., Wu, L., Zeng, L., 2018. An iterative local updating ensemble smoother for estimation and uncertainty assessment of hydrologic model parameters with multimodal distributions. Water Resour. Res. 54 (3), 1716–1733. https://doi.org/10.1002/2017WR020906.

Zhang, J., Zeng, L., Chen, C., Chen, D., Wu, L., 2015. Efficient Bayesian experimental design for contaminant source identification. Water Resour. Res. 51 (1), 576–598. https://doi.org/10.1002/2014wr015740.

Zhang, J., Zheng, Q., Wu, L., Zeng, L., 2020. Using deep learning to improve ensemble smoother: applications to subsurface characterization. Water Resour. Res. 56 (12). https://doi.org/10.1029/2020wr027399.

Zhang, T.-F., et al., 2019. Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks. Pet. Sci. 16 (3), 541–549. https://doi.org/10.1007/s12182-019-0328-4.

Zhou, B. et al., 2019. Semantic understanding of scenes through the ade20k dataset. Int. J. Comp. Vision, 127: 302-321, 10.48550/arXiv.1608.05442.

Zhu, J., Yeh, T.-C.J., 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. Water Resour. Res. 41 (7). https://doi.org/10.1029/2004WR003790.

Zhu, L., Dai, Z., Gong, H., Gable, C., Teatini, P., 2015. Statistic inversion of multi-zone transition probability models for aquifer characterization in alluvial fans. Stoch. Env. Res. Risk A. 30 (3), 1005–1016. https://doi.org/10.1007/s00477-015-1089-2.

Zhu, L., et al., 2020. The 3-D facies and geomechanical modeling of land subsidence in the Chaobai Plain, Beijing. Water Resour. Res. 56 (3). https://doi.org/10.1029/2019wr027026.

Zhu, L., Gong, H., Dai, Z., Guo, G., Teatini, P., 2017. Modeling 3-D permeability distribution in alluvial fans using facies architecture and geophysical acquisitions. Hydrol. Earth Syst. Sci. 21 (2), 721–733. https://doi.org/10.5194/hess-21-721-2017.