Petroleum Science 23 (2026) 1754–1772

Contents lists available at ScienceDirect

Petroleum Science

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

Original Paper

An interpretable attention-guided generative adversarial network framework with dual-domain learning for multi-condition constrained sedimentary facies modeling

Lei Liu [a,b,c,d], Wei Li [a,b,c], Jian Gao [e], Da-Li Yue [a,b,c,d,*], De-Gang Wu [a,b,c,d], Wu-Rong Wang [a,b,c], Jin Lin [a,b,c], Zhi-Bo Li [a,b,c,d], Qian Zhong [a,b], Jia-Gen Hou [b,c,d]

[a] Hainan Institute of China University of Petroleum, Sanya, 572025, Hainan, China
[b] State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[c] College of Geosciences, China University of Petroleum, Beijing, 102249, China
[d] College of Artificial Intelligence, China University of Petroleum, Beijing, 102249, China
[e] CNPC Science and Technology Research Institute Co., Ltd., Beijing, 100083, China

ARTICLE INFO

Article history:
Received 15 July 2025
Received in revised form
23 November 2025
Accepted 23 February 2026
Available online 26 February 2026

Edited by Xiu-Fang Hu

Keywords:
Sedimentary facies models
Attention-guided generative adversarial network
Interpretable framework
Sedimentary patterns
Multi-condition modeling

ABSTRACT

Sedimentary facies modeling is a critical approach for understanding geological phenomena, yet the strong heterogeneity of reservoir systems poses a serious challenge for their refined characterization. In this study, we innovatively propose an interpretable attention-guided generative adversarial network framework with dual-domain learning, which achieves precise sedimentary facies modeling under the constraints of well facies and soft probability data. Specifically, we first effectively extract and preserve prior information of sedimentary facies models from both spatial and frequency domain perspectives. Then, during simulation, to enhance the capability of the network model for finely characterizing complex heterogeneous models, cross-spatial attention mechanisms are designed to effectively capture short-range and long-range dependencies between multi-scale pattern features. Additionally, through systematic feature map visualization analysis, we elucidate the processes of conditional fitting and complex sedimentary facies model reconstruction, intuitively demonstrating the functional mechanisms of each module. Finally, systematic experiments are conducted on multiple datasets to validate the effectiveness of the proposed method. The results demonstrate that the generated sedimentary facies models exhibit high consistency with training datasets in terms of visual realism and statistical indicators. Quantitative comparisons reveal remarkable performance of the method, achieving low Wasserstein distance (0.09), Kernel Inception Distance (0.0017) and Kernel Maximum Mean Discrepancy (0.21). These findings further confirm the high realism of the generated realizations regarding pattern features. This study offers a reliable and practical method for geological reservoir modeling, thereby advancing quantitative, precise geological research with broad application prospects.

1. Introduction

Sedimentary facies models can reflect depositional environments and sedimentary processes of geological reservoirs, and their accurate characterization is of great importance in hydrocarbon exploration and development, carbon dioxide geological sequestration, and subsurface water exploitation (Chen et al., 2023; Liu and Liu, 2022; Pawar et al., 2016; Wang et al., 2024). Significant research efforts have been devoted to geostatistical modeling, and these methods are widely used in the simulation and assessment of subsurface sedimentary facies distributions (Jona Lasinio, 2015; Rezaee et al., 2013; Zakeri and Mariethoz, 2021). However, the reliance of these traditional geostatistical methods on the assumption of spatial stationarity poses considerable limitations when applied to highly heterogeneous subsurface sedimentary facies models, where this assumption is often

* Corresponding author.
E-mail address: yuedali@cup.edu.cn (D.-L. Yue).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.02.018

violated (Li et al., 2016; Liu et al., 2025; Zakeri and Mariethoz, 2021). These limitations manifest primarily in two key aspects: (1) Inadequate capture of non-stationary patterns and insufficient reconstruction of subsurface sedimentary facies models often lead to discontinuous simulation results (Fan et al., 2024a; Song et al., 2021a); (2) Constraints in statistical parameter estimation derived from observation data (e.g., wellbore and geophysical data) frequently introduce significant spatial uncertainty and representation biases (Cui et al., 2023; Jona Lasinio, 2015).

In recent years, data-driven deep learning techniques have provided novel solutions for precise geological characterization (Chi et al., 2025). Among these, generative adversarial networks (GANs) have made notable strides in subsurface sedimentary facies modeling (Feng et al., 2022; Cui et al., 2023, 2024a; Zhan et al., 2022; Zhang et al., 2019). By learning the latent spatial distribution patterns of training images, GANs can efficiently reproduce distinct and refined subsurface sedimentary patterns that align with the attribute distributions of the training data. Current research on GAN-based sedimentary facies modeling is directed towards two key improvements: first, optimizing feature extraction architectures to effectively capture complex geological features, and second, refining conditioning data fusion methods to yield more applicable outputs. Research in optimizing feature extraction architectures focuses on improving and applying GAN-derived models (Fan et al., 2024a; Hu et al., 2024; Liu et al., 2023; Zhang et al., 2022). Early studies typically employed fully convolutional neural networks, aiming to reproduce the spatial patterns of sedimentary facies models (Laloy et al., 2018; Liu et al., 2022; Ruffino et al., 2020). However, these models, due to the synchronous parameter updates across layers during training, often fail to adequately capture the variability in sedimentary features across different scales, thus limiting their simulation effectiveness. To address this challenge, researchers have adopted new simulation strategies, including multi-stage modeling (Fan et al., 2024b; Feng et al., 2025; Liu et al., 2023) and progressive simulation (Song et al., 2021a, 2022b; Tetteh et al., 2024). These strategies enable models to gradually transition from coarse to fine scales, thus facilitating the layer-by-layer learning of spatial geological features. Furthermore, the U-Net, with its encoder–decoder framework, has demonstrated significant performance in subsurface sedimentary facies modeling (Fan et al., 2024b). The encoder extracts deep features, while the decoder facilitates feature reconstruction. Skip connections between these components effectively mitigate information loss, further enhancing the reproduction of complex patterns. The static weights of traditional convolutions struggle to adapt to the dynamic variations in input data, and this fundamental limitation markedly hampers their effectiveness in dynamic feature extraction. Therefore, an increasing number of studies have integrated attention mechanisms, such as spatial attention (Zhang et al., 2024), channel attention (Liu et al., 2023), and coordinate attention (Liu et al., 2025), into the simulation networks. This integration enhances the flexibility and efficiency of network models in processing geological information, greatly improving their adaptability and performance in modeling (Cui et al., 2024a; Song et al., 2022b). Overall, the methods for improving feature extraction architectures are diverse. These optimization strategies effectively enhance the ability of the network models to characterize complex sedimentary facies models. However, these studies focus on the spatial-domain distribution patterns of geological features, while neglecting the rich detailed information in the frequency domain. In addition, subsurface sedimentary facies models exhibit distinct hierarchies, with complex spatial coupling relationships among sedimentary facies characteristics at varying scales. Without a reasonable mechanism to guide the capture of long-range dependency features in multi-scale sedimentary facies, network models struggle to effectively characterize and reproduce heterogeneous patterns.

When conducting multi-source conditioning data fusion, the U-Net architecture has demonstrated notable advantages (Pan et al., 2021; Zhang et al., 2021). The architecture integrates conditioning data into the encoding and decoding processes for efficient information fusion. Coupled with the iterative optimization of conditional loss functions, this approach substantially enhances the accuracy of conditional sedimentary facies models. To address complex application scenarios, researchers have adopted multi-stage input strategies, such as downsampling (Cui et al., 2024a; Song et al., 2021b, 2022b; Tetteh et al., 2024) and implicit encoding (Cui et al., 2024b). These strategies effectively capture spatial distribution relationships between multi-scale sedimentary facies features and conditioning data. Furthermore, to preserve spatial layout characteristics and propagate conditional semantics accurately, recent studies have innovatively introduced spatially-adaptive conditioning modules (Abdellatif et al., 2023; Liu et al., 2025; Sun et al., 2023). These modules enable GANs to implicitly learn the spatial distribution patterns of conditioning data and their complex associations with sedimentary facies models, thereby generating sedimentary facies models that are adaptively fitted to the conditioning data. Overall, the essence of conditional fusion research lies in elucidating the spatial distribution patterns of conditioning data and their spatial correspondence with the generated sedimentary facies models. However, the introduction of additional conditional loss terms significantly impacts the spatial pattern representations and training stability of GANs, making it challenging to balance pattern reproduction accuracy with conditional alignment. Furthermore, the data-driven conditional sedimentary facies modeling methods often exhibit “black box” characteristics, lacking interpretability in the simulation process. This can lead to unreliable results, thus limiting their credibility in practical engineering applications.

To address these aforementioned challenges, we develop an interpretable attention-guided GAN framework integrated with dual-domain learning. This framework facilitates the multi-condition controlled sedimentary facies modeling by comprehensively capturing data characteristics in both the spatial and frequency domains. Leveraging multiple attention mechanisms, it guides the simulation to focus on key patterns across multi-scale sedimentary facies. Furthermore, efficient and stable loss functions are introduced to generate realistic sedimentary facies models. The main innovations and contributions of this work are as follows:

- Develop a dual-domain learning module to simultaneously extract features from the spatial and frequency domains, leveraging their synergistic complementarity to achieve the accurate reconstruction of complex sedimentary facies models.
- Propose guided cross-spatial attention mechanisms to significantly enhance the characterization of critical sedimentary patterns. This enables the fusion of low-level conditioning features with high-level geological representations and employs a pairwise interaction mechanism to process pixel-level features.
- Design novel hybrid loss functions that combine gradient penalty, hinge loss, and conditional loss, and employ a weighted mechanism to balance these components to stabilize the training process of the model.
- Visualize the feature maps of different deep learning layers to intuitively demonstrate the simulation process of sedimentary

facies modeling, provide valuable insights into data-driven sedimentary facies modeling, and enhance the interpretability of the network model.
- Conduct a series of experiments using the conceptual river and the Brahmaputra River datasets, with validation results supported by comprehensive evaluation metrics, robustly demonstrating the superiority and reliability of our method in simulating complex sedimentary facies models.

2. Methodology

The two-dimensional discrete Fourier transform (2D-DFT) is employed to assist in constructing soft probability data. The GAN model primarily consists of a generator and a discriminator. This model guides key pattern information through residual spatial attention gates (RSAG) and pairwise-guided multi-scale spatial attention (PMSA), while simultaneously utilizing a dual-domain learning module (DDLB) to extract conditioning features. The hybrid loss functions are applied to stabilize the training process and facilitate conditional modeling. The comprehensive modeling framework encompasses the entire workflow, including data processing, training optimization, performance evaluation, and interpretability analysis.

2.1. Two-dimensional discrete Fourier transform (2D-DFT)

The discrete Fourier transform (DFT) is a widely used mathematical tool in image processing, enabling the transformation of images from the spatial to the frequency domain. In this study, 2D-DFT is employed to decompose images into the frequency domain and subsequently reconstruct the images (Amidror, 2013).

2.1.1. Theoretical foundation

For an image $f(h,w)$ with a resolution of $H \times W$, where $h = 0, 1, 2, \cdots, H-1$ denotes the row index and $w = 0, 1, 2, \cdots, W-1$ denotes the column index, the 2D-DFT is defined as:

$$F_{\mathrm{fre}}(u,v) = \sum_{h=0}^{H-1}\sum_{w=0}^{W-1} f(h,w)\exp\left[-2\pi i\left(\frac{uh}{H}+\frac{vw}{W}\right)\right] \tag{1}$$

where $F_{\mathrm{fre}}(u,v)$ represents the frequency domain representation, with $u = 0,1,2,\cdots,H-1$, and $v = 0, 1, 2, \cdots, W-1$ being the vertical and horizontal frequency indices, respectively; $\exp(\cdot)$ denotes the exponential function with base e; i represents the imaginary unit.

2.1.2. Feature extraction and inverse transform

In the frequency domain, after applying a frequency shift operation to center the zero-frequency component, low-frequency components correspond to smooth regions of the image, while high-frequency components capture details such as edges. A circular low-pass filter mask $F_{\mathrm{mask}} \in \mathbb{R}^{H\times W\times 1}$ is designed with its center at $(H/2, W/2)$, defined as:

$$F_{\mathrm{mask}} = \begin{cases} 1, if \left(u-\frac{H}{2}\right)^2+\left(v-\frac{W}{2}\right)^2 \le R^2 \\ 0, \text{otherwise} \end{cases} \tag{2}$$

where $R \in \left[\max\left(\frac{H}{32},\frac{W}{32}\right), \min\left(\frac{H}{8},\frac{W}{8}\right)\right]$ serves as the radius for extracting data within a specific range, and this range is chosen to balance the trade-off between resolution and data quality based on empirical observations (Fig. 1). The low-frequency components are extracted by applying the mask (F_{mask}), expressed as:

$$F_{\mathrm{low}}(u,v) = F_{\mathrm{fre}}(u,v)\cdot F_{\mathrm{mask}}(u,v) \tag{3}$$

Using the extracted low-frequency data $F_{\mathrm{low}}(u, v)$, the low-frequency image is reconstructed via the inverse 2D-DFT:

$$f_{\mathrm{low}}(h,w) = \frac{1}{HW}\sum_{u=0}^{H-1}\sum_{v=0}^{W-1} F_{\mathrm{low}}(u,v)\exp\left[2\pi i\left(\frac{uh}{H}+\frac{vw}{W}\right)\right] \tag{4}$$

As shown in Fig. 1, the increase in the radius of the circle gradually expands the spectral coverage. Specifically, when the radius is small, the spectral energy is primarily concentrated in the low-frequency region, and the reconstructed image only reflects the overall outline of the sedimentary facies model. As the radius increases, high-frequency components become more prominent, and the detailed features of the sedimentary facies models in the reconstructed image gradually become clearer.

2.2. The architecture of the network

2.2.1. Cross-spatial attention mechanisms with guidance

The fusion process of the attention modules is shown in Fig. S1a of the Supporting material.

(1) Residual spatial attention gate (RSAG)

Conditioning data inherently contain rich prior knowledge, enabling precise extraction of key layout and distribution features, which are crucial for conditional sedimentary facies modeling. The high-level geological features in the simulation can effectively guide this process. To more comprehensively capture key spatial features, a residual spatial attention gate (RSAG) was proposed (Fig. 2). This gate unit is capable of performing cascaded fusion of low-level conditioning features and high-level geological representations, effectively enhancing the model's perception of key information without significantly increasing its complexity (Liu et al., 2025; Oktay et al., 2018).

Specifically, the low-level conditional features ($\overline{F}_{\mathrm{con}}$) and high-level geological representations ($\overline{F}_{\mathrm{geo}}$) are each processed through two consecutive convolutional mapping transformations. Subsequently, by performing element-wise addition, the spatial layout details of low-level features are effectively integrated with high-level features. Following this, the application of a nonlinear mapping via the activation function (θ_1) enhances the model's representation capacity. To further exploit internal information within the features and the implicit relationships between the conditional features and geological representations, an additional convolutional transformation is incorporated. Finally, the result is mapped to an attention feature map (Q_{map}) through the activation function (σ), with values ranging from 0 to 1, representing the attention weight for each spatial location. The process can be defined as:

$$Q_{\mathrm{map}} = \sigma\left(\psi_3\left[\theta_1\left(\psi_1\left[\overline{F}_{\mathrm{con}}\right]+\psi_2\left[\overline{F}_{\mathrm{geo}}\right]\right)\right]\right) \tag{5}$$

where ψ_1 and ψ_2 represent 1×1 convolutions followed by batch normalization (BN) layers, ψ_3 denotes a 1×1 convolution followed by a BN layer, θ_1 is a ReLU function, and σ is a sigmoid function.

During the simulation, this attention map can implicitly generate soft region proposals, thereby effectively highlighting key features closely associated with conditional sedimentary facies modeling. To further enhance the generalization capability and stability of conditional simulation, the module is optimized with a residual structure, and its computational process can be expressed as follows:

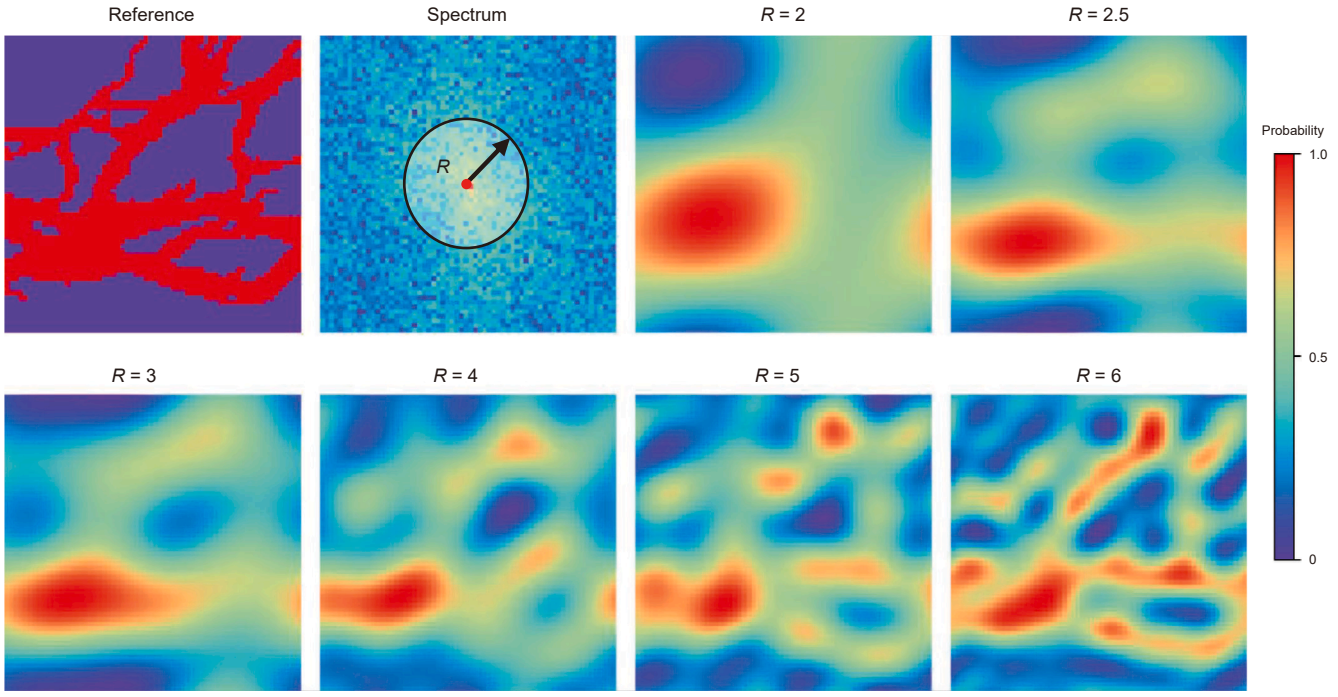

Fig. 1. The feature extraction and inverse transform process of the 2D-DFT. The circular mask in the spectrum is used to extract key spectral features from the Fourier transform domain. The values of H and W are both 64 in this case. The parameter R represents the radius of the reconstruction region, with larger values of R corresponding to higher-resolution reconstructions in the inverse transform.

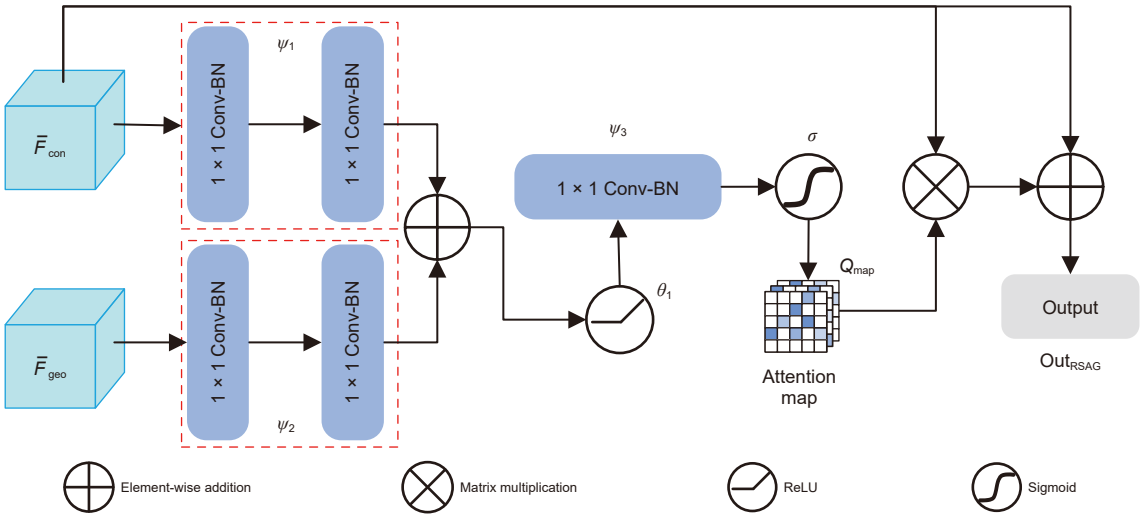

Fig. 2. Architecture of RSAG.

$$Out_{RSAG} = (Q_{map} + 1) \times \bar{F}_{con} \quad (6)$$

(2) Pairwise-guided multi-scale spatial attention (PMSA)

In sedimentary facies modeling, GANs often fail to effectively capture coexisting multi-scale geological information. While introducing traditional spatial attention mechanisms can enhance the network's focus on critical information and reduce computational costs, this approach may compromise the integrity of feature capture, potentially losing fine-scale details. In contrast, retaining complete information in each feature channel and further capturing pairwise long-range dependencies between pixels can enable more efficient reconstruction of structural patterns across different scales (Liu et al., 2025; Ouyang et al., 2023).

Therefore, we propose the pairwise-guided multi-scale spatial attention (PMSA). PMSA integrates feature information across different scales by leveraging positional learning and spatial aggregation, emphasizing key feature details (Fig. 3). Multi-scale geological features (e.g., $\tilde{F}_1$, $\tilde{F}_2$, $\tilde{F}_3$, and $\tilde{F}_4$) from various simulation stages are first concatenated along the channel dimension to form the feature ($\hat{F}_{concat} = [\hat{F}_0, \hat{F}_g, ..., \hat{F}_{\tilde{G}-1}]$). Subsequently, this feature is divided into G feature groups along the channel dimension. For each feature group ($\hat{F}_g \in \mathbb{R}^{\tilde{C}//\tilde{G}\times\overline{H}\times\overline{W}}$), three defined features ($\hat{F}_{aggr}$, $\hat{F}_{Posit1}$, and $\hat{F}_{Posit2}$) are introduced to perform two

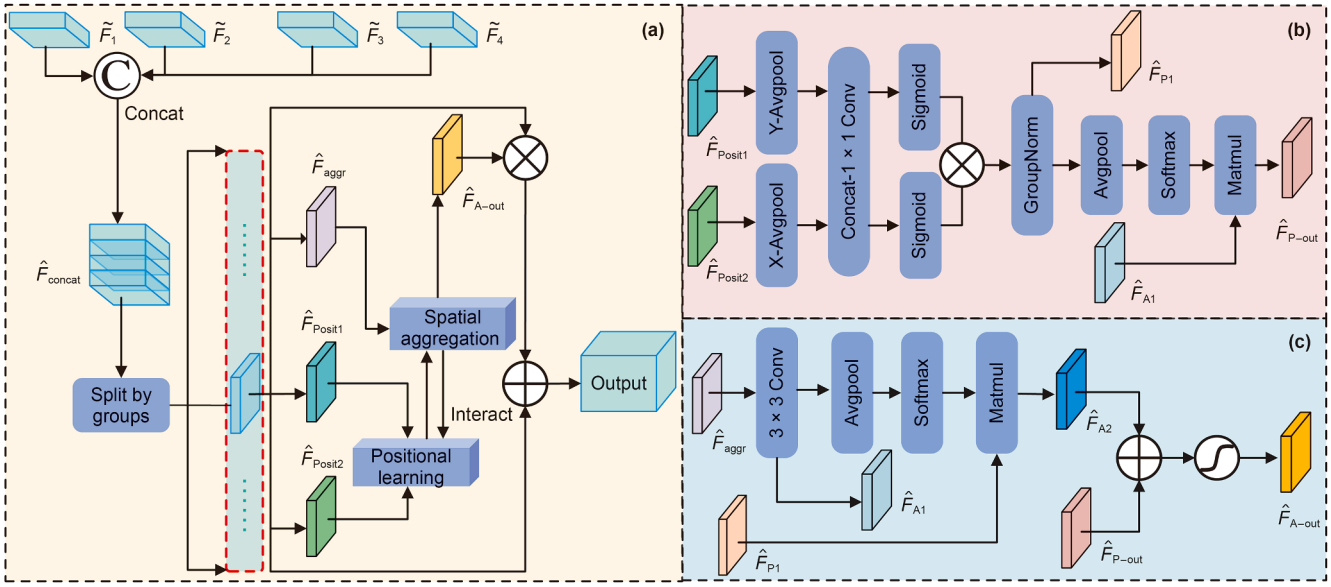

Fig. 3. **(a)** Architecture of PMSA; **(b)** positional learning; **(c)** spatial aggregation.

parallel cross-spatial branch tasks: key positional feature learning and effective spatial information aggregation (Fig. 3(a)). Among these, $\widehat{F}_{\text{Posit1}}$ and $\widehat{F}_{\text{Posit2}}$ are employed for pixel-level positional information learning and to execute various transformation operations in the X and Y directions (Fig. 3(b)), such as pooling in different directions, convolutional operations, and activation function mapping, to highlight spatial information most relevant to the simulation results, thereby generating the feature matrix ($\widehat{F}_{\text{P-out}}$). Specifically, the features of the cl^{th} channel after pooling at $\overline{H}$ and $\overline{W}$ are represented as:

$$E_{cl}{}^{\overline{H}}(\overline{H}) = \frac{1}{\overline{W}} \sum_{0 \le j \le \overline{W}} \widehat{f}_{cl}(\overline{H}, j) \tag{7}$$

$$E_{cl}{}^{\overline{W}}(\overline{W}) = \frac{1}{\overline{H}} \sum_{0 \le k \le \overline{H}} \widehat{f}_{cl}(k, \overline{W}) \tag{8}$$

where $\widehat{f}_{cl}$ indicates the input features of cl^{th} channel, $E_{cl}{}^{\overline{H}}(\overline{H})$ and $E_{cl}{}^{\overline{W}}(\overline{W})$ denote the corresponding outputs.

As shown in Fig. 3(c), $\widehat{F}_{\text{aggr}}$ primarily captures short-range and long-range key pattern information through pixel-level pairwise interactions between cross-spatial features ($\widehat{F}_{\text{P1}}$, $\widehat{F}_{\text{A1}}$, $\widehat{F}_{\text{A2}}$, and $\widehat{F}_{\text{P-out}}$), generating corresponding attention weights ($\widehat{F}_{\text{A-out}}$). The attention weight matrix contains multi-scale feature interaction information. In the module design, we also incorporated a residual structure to enhance the training stability and generalization performance of the sedimentary facies modeling.

2.2.2. Generator architecture

The proposed generator employs an efficient encoder–decoder architecture designed to generate distinct and refined representations of sedimentary facies models. As shown in Fig. 4(a), the inputs to the generator include conditioning data and the random latent vectors ($\mathbf{z}$) from a standard normal distribution. The backbone of the generator architecture is jointly composed of four Encoder Blocks, four Decoder Blocks, and the corresponding UpBlocks. The encoding stage is mainly responsible for extracting key pattern features from the conditioning data layer by layer, progressively constructing high-level feature representations. The decoding stage converts the features into distinct sedimentary facies models through feature fusion and gradual decoding.

Within the encoding path, each Encoder Block adopts a residual structure (Fig. 4(b)). This design enhances model stability by progressively extracting features. Notably, during the downsampling, high-resolution geological features are compressed into low-resolution geological representations, which is typically accompanied by the loss of fine information. The conventional downsampling approaches predominantly focus on spatial domain transformations, often overlooking the rich information embedded in the frequency domain. To effectively preserve geological details and conditional representations, we introduce a novel dual-domain learning block (DDLB). This block employs parallel downsampling in both spatial and frequency domains to gradually reduce spatial dimensions while retaining critical features through complementary dual-domain interactions (Fig. 4(c)).

Specifically, the spatial domain branch utilizes Maxpool and Avgpool to downsample and extract the overall conditional distribution, followed by parallel 3 × 3 convolutions to capture diverse features. These are further refined using a 3 × 3 convolution, BN layer and ReLU activation function to extract meaningful spatial domain feature representations. In contrast, the frequency domain downsampling transforms input data into the frequency domain using Haar wavelet transform (Xu et al., 2023). The spatial resolution of the two-dimensional image will be reduced to half after applying the Haar wavelet transform (as detailed in Text S1 of Supporting material). This process simultaneously generates four distinct components: the low-frequency component (*M_A*), along with the horizontal (*M_H*), vertical (*M_V*), and diagonal (*M_D*) components that characterize high-frequency information. These component features reflect the characteristic patterns of sedimentary facies models, and their concatenation along the channel dimension collectively reconstructs the complete feature pattern (Fig. 4(b)). Subsequently, a 3 × 3 convolution, BN layer, and ReLU activation function are applied to produce the frequency-domain output. Ultimately, by integrating the output results from the spatial and frequency domains, cross-domain information

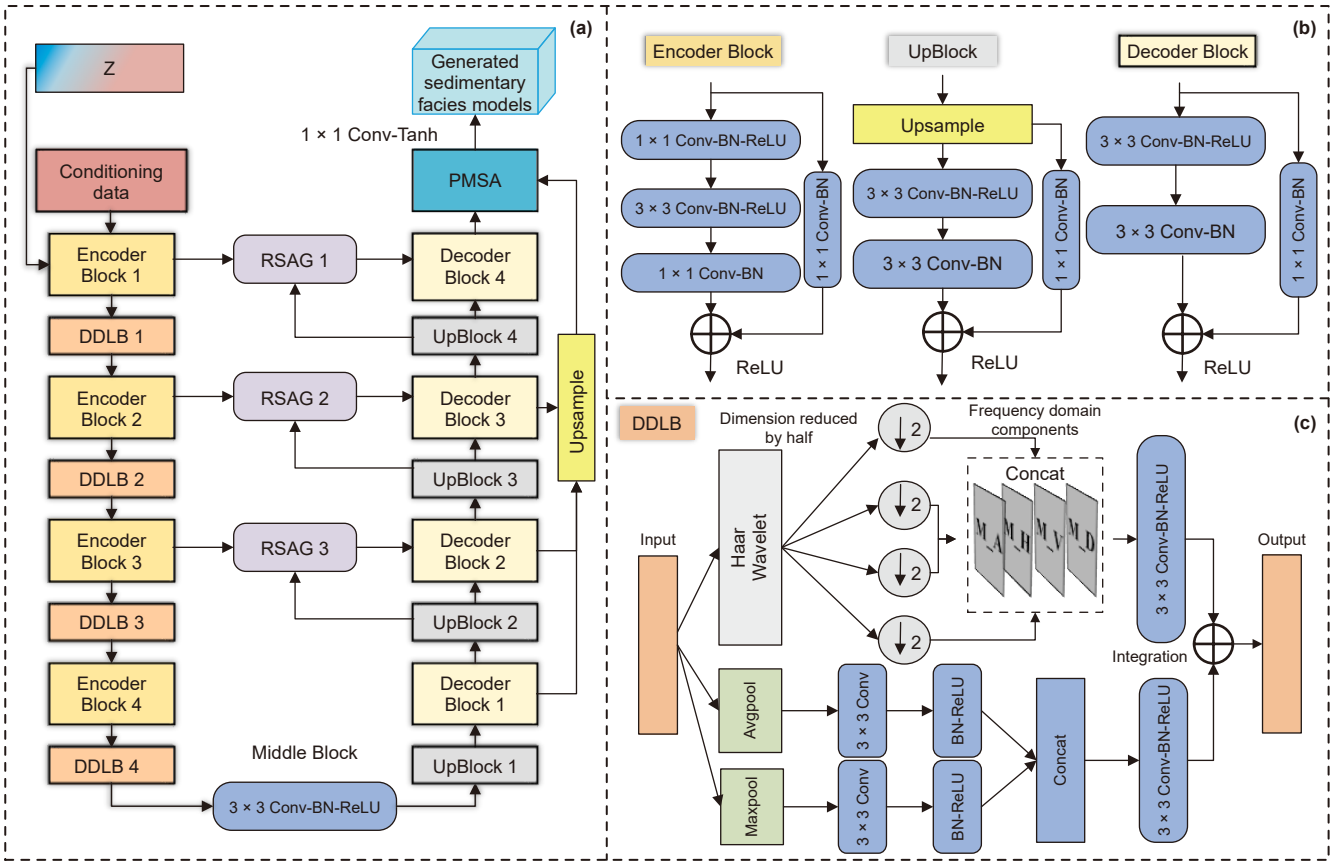

Fig. 4. **(a)** The architecture of generator; **(b)** details of Encoder Block, UpBlock, and Decoder Block; **(c)** dual-domain learning block (DDLB). 3 × 3 Conv-BN-ReLU indicates a 3 × 3 convolution, BN layer, and ReLU activation.

complementarity is achieved, significantly enhancing the model's performance in detail extraction and robustness. A schematic diagram of the fusion process for DDLB is provided in Fig. S1b of the Supporting material.

To enhance information flow while minimizing the loss of critical information, RSAGs are deployed between the encoder and decoder. The RSAGs extract features from the encoder blocks and fuse them with the corresponding Decoder Block features, generating regional soft proposals to emphasize features highly relevant to conditional simulation. Along the decoding path, features progressively restore spatial resolution through four UpBlocks and Decoder Blocks. Each UpBlock employs the structure depicted in Fig. 4(b) to preliminarily integrate feature information and upscale the feature maps. The Decoder Blocks further refine resolution. Additionally, after the fourth Decoder Block, a PMSA is incorporated. This block globally aggregates features from different stages, achieving cross-scale integration to produce realistic sedimentary facies models. Finally, the fused features are mapped into detailed sedimentary facies models via a 1 × 1 convolution followed by the Tanh activation function.

2.2.3. Discriminator architecture

The proposed discriminator is primarily based on the PatchGAN architecture, designed to evaluate the differences between generated sedimentary facies models and true sedimentary facies models, with its architecture illustrated in Fig. 5. The backbone of the discriminator comprises multiple convolutional layers, specifically five convolutional blocks. Each block employs a 4 × 4 convolution and integrates the LeakyReLU activation function to enhance the nonlinear feature extraction. To improve the performance of the discriminator, the model incorporates PMSA. By fusing multi-scale features extracted from the outputs of the first to fourth convolutional blocks, PMSA captures pixel-level discrepancies between images from different sources via attention-based weighting, significantly improving the model's sensitivity to geological details. Furthermore, the mini-batch standard deviation (MSD) technique is employed to enhance the ability of the discriminator to perceive the diversity of generated realizations (Song et al., 2021a). This technique captures statistical variations within mini-batches, thereby encouraging the generator to produce more realistic and diverse sedimentary facies models.

2.3. Hybrid loss functions

The loss function is a crucial component in sedimentary facies modeling based on conditional GANs, primarily encompassing two categories: conditional loss and adversarial loss. Conditional loss aims to ensure the generated realizations align with limited observation data, while adversarial loss serves to simulate the sedimentary facies models similar to the training images. The inputs to the generator include the latent vectors ($\mathbf{z}$) and conditioning data (C), which comprise hard well facies data (C_{hard}) and soft probability data (C_{soft}). Hard conditional loss is typically derived by directly calculating the discrepancy between the generated realizations ($G(\mathbf{z}, C)$) and the actual geological references ($\overline{X}$) at observation locations. Under the constraints of hard

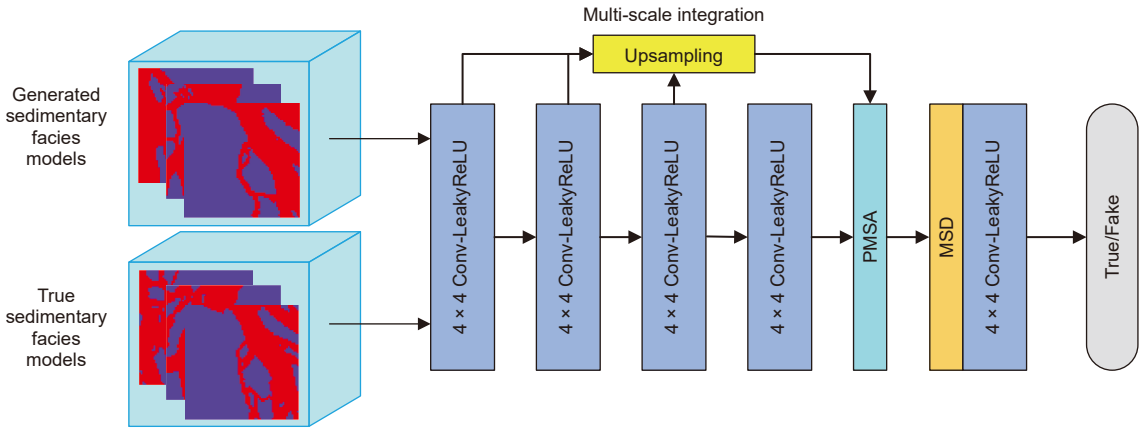

Fig. 5. The architecture of discriminator.

conditional loss, the simulation realizations can be ensured to match the well facies data. The loss can be expressed as:

$$L_{\text{con\_hard}} = \|C_{\text{mask}} \odot (G(\mathbf{z}, C)) - C_{\text{hard}}\|_2 \tag{9}$$

where C_{mask} represents the mask of well locations, with well locations set to 1 and other locations set to 0. $\odot$ denotes the element-wise product. $\|\cdot\|_2$ is the L2-norm (Euclidean norm).

The acquisition of soft probability data typically requires an additional transformation (S_{soft}), which will be elaborated in Section 2.4.1. Soft conditional loss can facilitate models in achieving reasonable distribution constraints under conditions of uncertainty. A similar approach to compute the soft probability data discrepancy between the transformed generated realizations $S_{\text{soft}}(G(\mathbf{z}, C))$ and C_{soft} is defined as follows:

$$L_{\text{con\_soft}} = \left\| S_{\text{soft}}(G(\mathbf{z}, C)) - C_{\text{soft}} \right\|_2 \tag{10}$$

In sedimentary facies modeling, the adversarial loss function constrains the GANs to ensure the realizations align with the spatial structure and patterns of the references. The generator should simulate high-order geological expressions, while the discriminator is tasked with distinguishing between $G(\mathbf{z}, C)$ and $\overline{X}$. However, imbalanced training during the simulation process can lead to modeling failures, particularly when the discriminator's performance is too strong, causing the generator's gradients to vanish. To achieve a balanced competition between the generator and the discriminator, enabling the generator to be continuously optimized, non-saturating hinge loss functions were designed to stabilize the adversarial training between the generator and the discriminator. It can be formulated as:

$$\begin{cases} L(D) = \mathrm{E}_{\mathbf{z}\sim P_{\mathbf{z}}, \overline{X}\sim P_{\overline{X}}}[\max(0, 1 + D(G(\mathbf{z}))) + \max(0, 1 - D(\overline{X}))] \\ L(G) = -\mathrm{E}_{\mathbf{z}\sim P_{\mathbf{z}}} D(G(\mathbf{z})) \end{cases} \tag{11}$$

where G and D denote the generator and discriminator, respectively, while $L(G)$ and $L(D)$ represent their initial loss functions. This formula strategically balances the antagonism between the discriminator and the generator with a clear boundary, ensuring that even when the discriminator performs strongly, the corresponding linear output $L(G)$ retains a gradient. Furthermore, to meet the requirements of conditional fitting and pattern reproduction in sedimentary facies modeling, we designed hybrid loss functions:

$$\begin{cases} L(D)_{\widehat{\text{hybrid}}} = \mathrm{E}_{\mathbf{z}\sim P_{\mathbf{z}}, \overline{X}\sim P_{\overline{X}}, C_{\text{hard}}\sim P_{C_{\text{hard}}}, C_{\text{soft}}\sim P_{C_{\text{soft}}}} \Big[\max\Big(0, 1 + D\Big(G\Big(\mathbf{z}, C_{\text{hard}}, C_{\text{soft}}\Big)\Big)\Big) + \max(0, 1 - D(\overline{X}))\Big] \\ L(G)_{\text{hybrid}} = -\mathrm{E}_{\mathbf{z}\sim P_{\mathbf{z}}, C_{\text{hard}}\sim P_{C_{\text{hard}}}, C_{\text{soft}}\sim P_{C_{\text{soft}}}} D\Big(G\Big(\mathbf{z}, C_{\text{hard}}, C_{\text{soft}}\Big)\Big) \\ +\lambda_1 L_{\text{con\_hard}} + \lambda_2 L_{\text{con\_soft}} \end{cases} \tag{12}$$

where λ_1 and λ_2 are weight parameters used to balance well facies and probability constraints. Additionally, to enhance training stability, we make the discriminator satisfy the 1-Lipschitz constraint, thus adding a gradient penalty term to $L(D)_{\widehat{\text{hybrid}}}$ (Gulrajani et al., 2017; Liu et al., 2025). The formula is as follows:

$$L(D)_{\text{hybrid}} = L(D)_{\widehat{\text{hybrid}}} + \lambda_{\text{gp}} \mathrm{E}_{\widehat{X}\sim P_{\widehat{X}}} \left[\left(\|\nabla_{\widehat{X}} D(\widehat{X})\|_2 - 1 \right)^2 \right] \tag{13}$$

where $\widehat{X}$ is defined as $\alpha\overline{X} + (1 - \alpha) G\Big(\mathbf{z}, C_{\text{hard}}, C_{\text{soft}}\Big)$, α is sampled from Uniform $[0, 1]$ and λ_{gp} is the gradient penalty coefficient.

The proposed hybrid loss functions can achieve multi-conditional simultaneous optimization. Through the non-saturating loss and the designed boundary, they effectively reduce mode collapse and training imbalance issues. The intuitive process of optimizing the hybrid loss functions can be found in Supporting material (Fig. S2).

2.4. Modeling framework of sedimentary facies models

The fundamental methodology of this paper has been outlined in the preceding sections, and this section systematically introduces the overall modeling framework (Fig. 6).

2.4.1. Dataset processing module

Samples were obtained through random rotation and cropping operations applied to the original dataset. By randomly sampling the training images, we obtained hard well facies data, with the number of conditioning points ranging between 8 and 16. Soft probability data primarily represents the distribution trends of the sedimentary facies models. However, it is worth noting that the quality of the actual soft probability data acquired is often influenced by factors such as heterogeneous reservoir distributions and noise introduced by the acquisition equipment. Accounting for these factors, we introduced uncertainty into the sedimentary

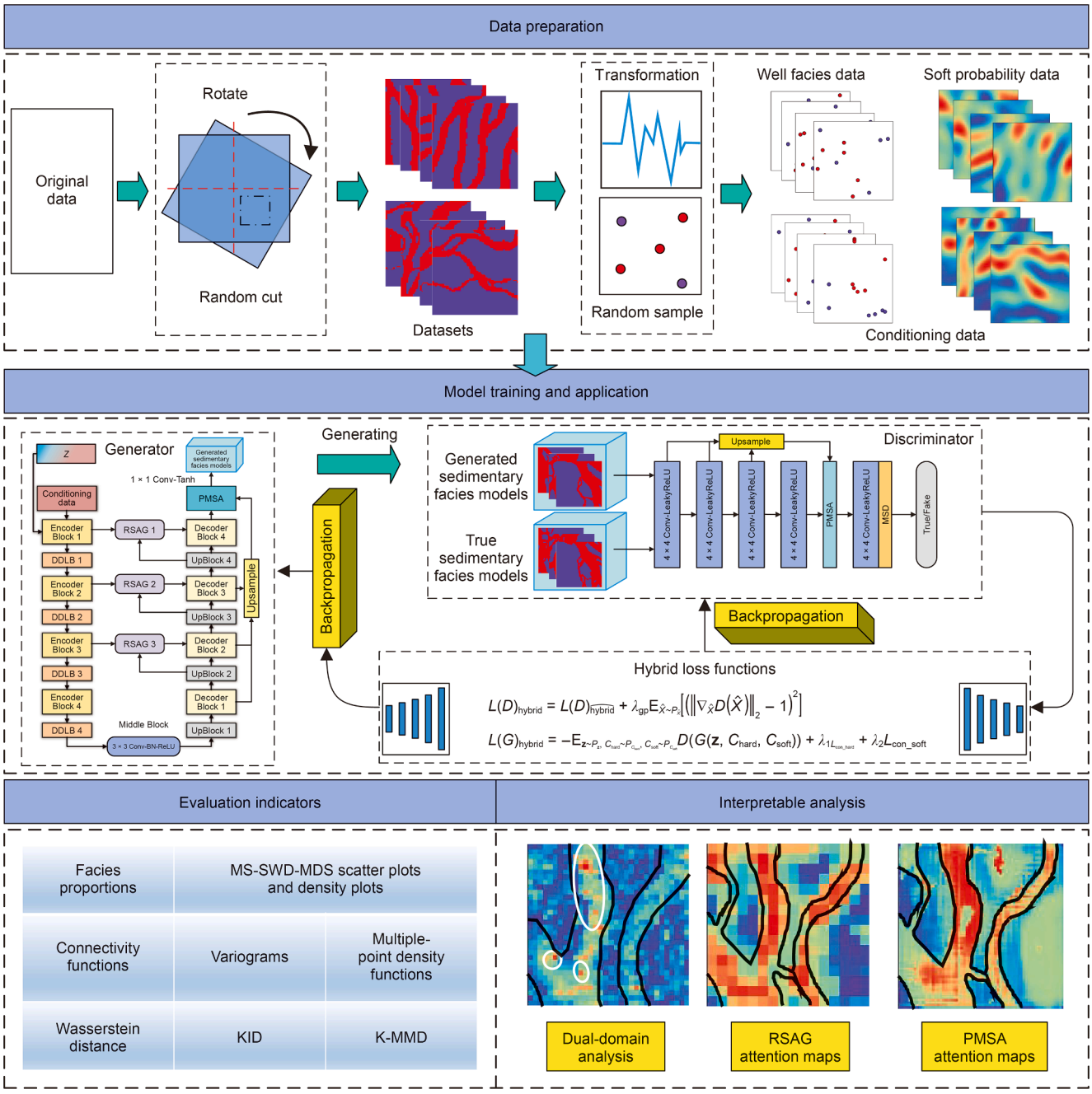

Fig. 6. Framework of our proposed method. It comprehensively covers the entire process of sedimentary facies modeling.

facies models by applying random morphological opening or closing operations ($\widehat{N}_{\mathrm{Morp}}$) and adding Gaussian filtering (N_{Noise}) to each sample ($f(h,w)$). The formula for this operation is as follows:

$$f_{\mathrm{Noise}}(h,w) = N_{\mathrm{Noise}}\left(\widehat{N}_{\mathrm{Morp}}[f(h,w)]\right) \tag{14}$$

Subsequently, we applied the 2D-DFT process, as described in Section 2.1.2, to $f_{\mathrm{Noise}}(h,w)$ to simulate geophysical transformations. By adjusting the value of R (ranging from 2 to 8) and utilizing the inverse transformation, we generated soft probability data at varying resolutions. For each dataset, 29,000 samples were allocated for training, and 1000 samples were reserved for testing.

2.4.2. Modeling procedure

During simulation, the generator takes conditioning data and random vectors as inputs, progressively learning high-order sedimentary patterns through multi-layer networks to accurately reproduce sedimentary facies models. The generated realizations and real sedimentary facies models are then fed into the discriminator, which distinguishes between them. The hybrid loss functions are employed to evaluate conditioning data consistency and pattern reproduction, and the backpropagation algorithm can optimize the generator and discriminator parameters separately to minimize the hybrid loss, enabling the generator to create more realistic and detailed sedimentary facies models (Fig. 6). The

adversarial interactions between them drive continuous optimization.

2.4.3. Evaluation indicators

This study accounts for key factors such as spatial distribution characteristics, pattern regularity, heterogeneity, and spatial continuity of sedimentary facies models. We employed comprehensive evaluation indicators, including facies proportions, multiscale sliced-Wasserstein distance and multidimensional scaling (MS-SWD-MDS) (Song et al., 2022a), connectivity functions, variograms and multi-point density functions, to validate the reliability of the simulation results and proposed method (Fig. 6). To further confirm the method's effectiveness, we conducted comparative analyses using commonly accepted similarity evaluation indicators in the field of computer vision (Fig. 6), including the Wasserstein distance, the Kernel Inception Distance (KID), which highlights perceptual quality, and the Kernel Maximum Mean Discrepancy (K-MMD), which assesses statistical divergence.

2.4.4. Interpretability analysis

Deep learning operates as an inherently opaque end-to-end process with limited intrinsic interpretability. To address this challenge, this study systematically analyzes the internal mechanisms of deep learning, thoroughly investigating the features learned during the simulation process and their significance. Specifically, the analysis is conducted across two dimensions: i) visualizing spatial and frequency domain analysis results to interpret key detailed features and the information they retain; ii) mapping attention mechanism weights to highlight salient features and their impact on simulation results. This multi-perspective analysis enhances the interpretability and reliability of sedimentary facies modeling.

3. Experiments

The modeling was performed with Python 3.9 and PyTorch 2.6.0 on a workstation equipped with an 11th Gen Intel® Core™ i7-11700K CPU at 3.6 GHz, 32 GB RAM, and an NVIDIA GeForce RTX 3070 Ti GPU. Tables S1 and S2 in the Supporting material provide the detailed parameter settings for the generator and discriminator. The GAN model was trained using a mini-batch gradient descent strategy with the Adam optimizer and mixed precision training.

3.1. Conceptual river

In this test, the GAN model was trained using the conceptual river dataset with 1250 × 1250 pixels (as detailed in Fig. S3a of Supporting material). Through random rotation and cropping operations on the original conceptual river, 30,000 samples were obtained, each with a resolution of 64 × 64 pixels (Fig. 6). The training data contains both channel and floodplain facies and shows typical diversity, primarily as continuously sinuous channels. Through multiple simulation experiments, we determined the appropriate weights ($\lambda_1 = 5$, $\lambda_2 = 0.5$ and $\lambda_{gp} = 10$) for the hybrid loss functions. We then obtained multiple geological realizations under the constraints of input hard well facies data and soft probability data (Fig. 7). The generated realizations are highly consistent with the corresponding conditioning data, while also exhibiting geological distribution characteristics and diversity that are remarkably similar to their references. The ninth column presents the frequency distribution of 300 realizations based on their corresponding conditioning data (Fig. 7). A comparison between the fourth and tenth columns reveals that the frequency density of multiple realizations consistently falls within the

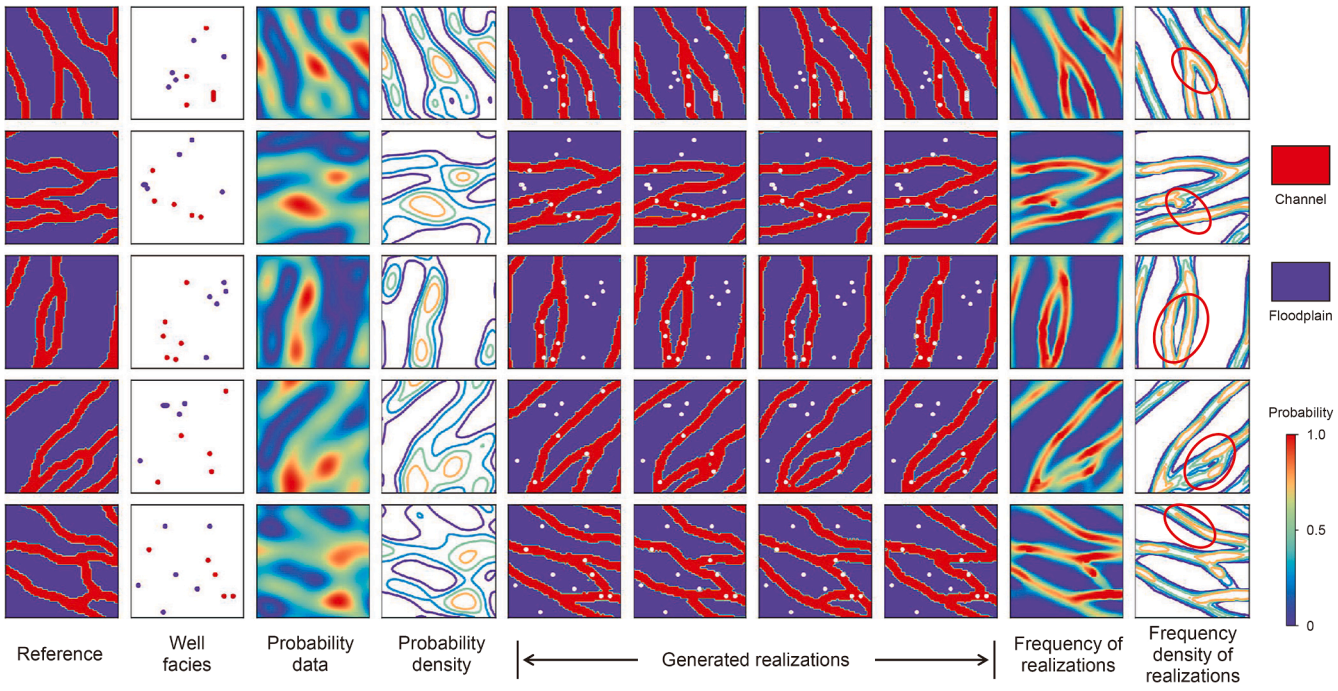

Fig. 7. Columns one to four display references, well facies data, probability maps, and probability density, respectively. Columns five to eight show the generated realizations, and the remaining columns show calculated frequency and frequency density maps derived from 300 realizations. The white scatter points indicate the well locations.

probability density of the soft probability data. Moreover, the frequency density plots successfully reproduce fine-scale geological features (marked by red circles in Fig. 7) absent in the input probability maps (e.g., connectivity of channel branches and channel sinuosity), indicating a low level of uncertainty in the realizations. This phenomenon can be explained from two perspectives: i) The simulation results are constrained not only by probability data but also by hard well facies data, and these constraints effectively narrow the range of possible realizations. ii) During the conditional simulation, the generator adeptly learns and captures patterns of spatial morphology in the complex sedimentary facies models. Consequently, the process achieves effective pattern fitting, which significantly compresses the uncertainty space of the simulation results.

Furthermore, we generated 100 corresponding realizations based on different conditioning data. The obtained geological realizations exhibit a high correlation with the references in terms of facies proportions (Fig. 8(a)). Notably, although facies proportions were not explicitly included as input conditions, the GAN model effectively learned the implicit features and associated attribute distributions during the simulation process. Additionally, these realizations correspond closely with the references in the MS-SWD-MDS scatter plots, and their MS-SWD-MDS density plots show a high degree of overlap (Fig. 8(b) and (c)). We also selected a specific set of conditioning data and generated 50 geological realizations based on it. Fig. 8(d)–(g) compares the spatial variability and continuity between these realizations and the corresponding reference, indicating that the blue curves of the realizations consistently fluctuate closely around the red curve of the reference. The statistical results confirm that the proposed simulation method effectively captures geological features and generates sedimentary facies models with patterns similar to the reference.

3.2. The Brahmaputra River

The heterogeneous Brahmaputra River dataset comprising 3273 × 2920 pixels (as detailed in Fig. S3b of Supporting material) was employed for the experiment. A total of 30,000 samples, each at a resolution of 64 × 64 pixels, were generated by implementing random rotation and cropping techniques. The training data contains both channel and floodplain facies. The Brahmaputra River dataset exhibits more complex multi-scale channel characteristics. Specifically, the dataset not only includes multi-level branching river systems but also demonstrates significantly higher heterogeneity in the spatial distribution patterns and morphological features, allowing us to explore the validity and flexibility of the simulation method. Through substantial training (Tables S3 and S4 of the Supporting material), suitable hyperparameters of hybrid loss functions ($\lambda_1 = 4$, $\lambda_2 = 0.5$ and $\lambda_{gp} = 10$) were selected. Constrained by the conditioning data, multiple realizations were constructed (Fig. 9). The generated realizations, presented in columns five to eight, demonstrate strong correspondence with both the well facies data and the soft probability data. The generated realizations exhibit distributional characteristics highly similar to the references while displaying a broader range of geological feature expressions, particularly in the distribution patterns of channel branching across different scales, clearly indicating sedimentary facies patterns consistent with the references. This diversity not only reflects the model's capability to capture complex geological structures but is particularly evident in the intricate and detailed features of channel branching and confluence, which further suggest a complex sedimentary dynamic environment. These observations, based on comparative analysis between the realizations and the references, validate the reliability of this method in reproducing certain sedimentary dynamics patterns.

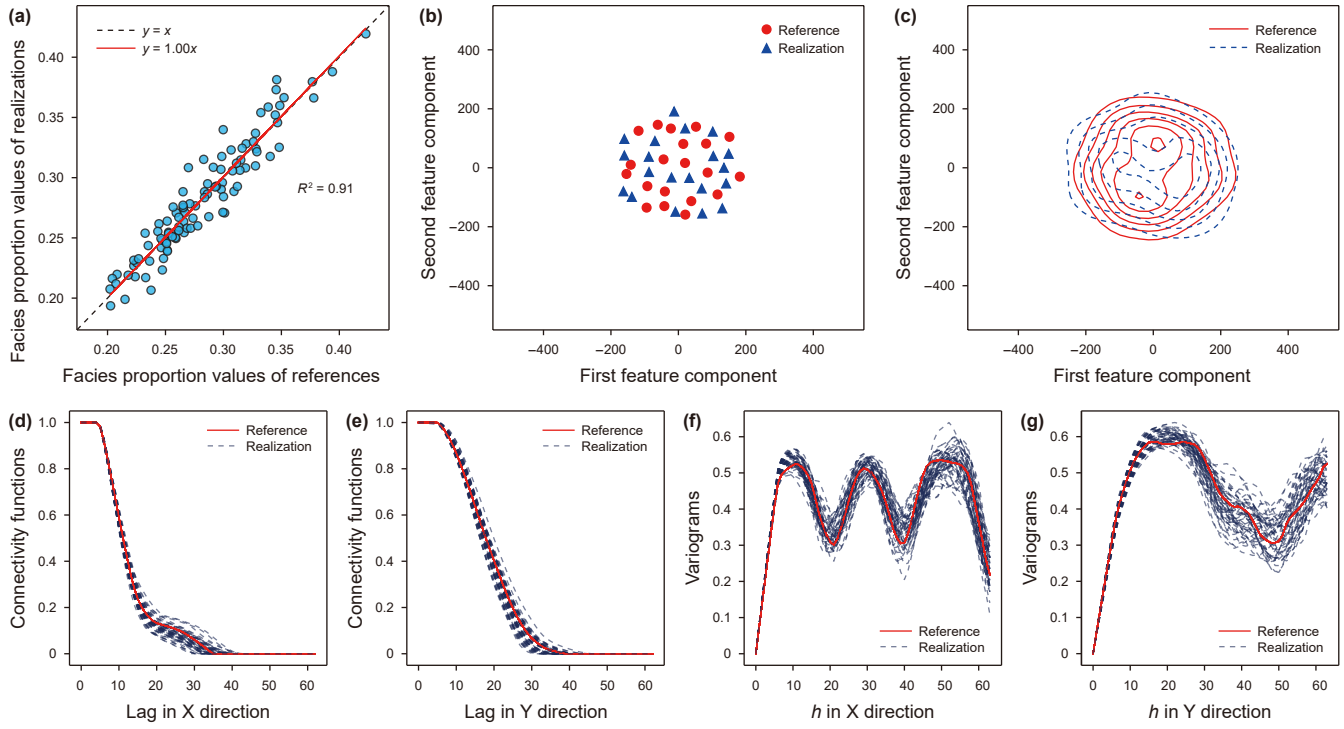

Fig. 8. Statistical comparison between realizations and references for conceptual river. **(a)** Facies proportions; **(b–c)** MS-SWD-MDS scatter plots and density plots; **(d–e)** connectivity functions in both X and Y directions; **(f–g)** variograms in both X and Y directions.

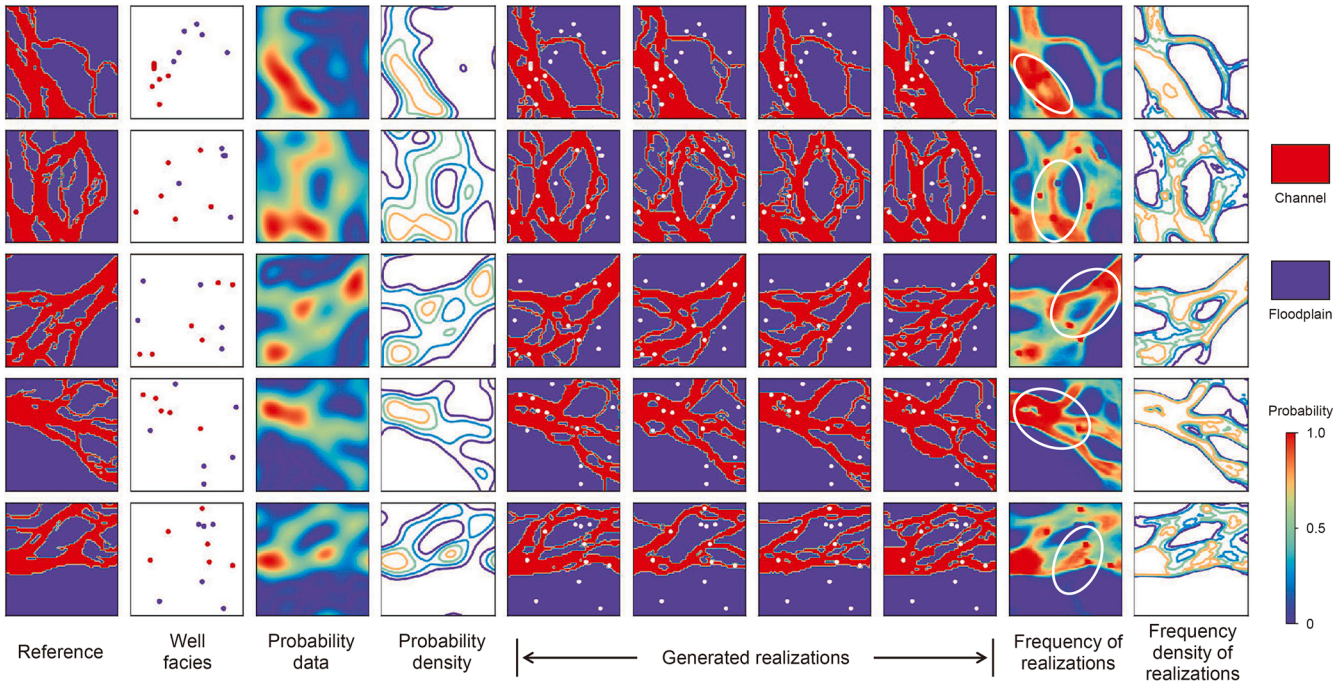

Fig. 9. Columns one to four display references, well facies data, probability maps, and probability density, respectively. Columns five to eight show the generated realizations, and the remaining columns show calculated frequency and calculated frequency density of realizations. The white scatter points indicate the well locations.

Compared to the probability maps, the frequency distribution of the 300 generated realizations using conditioning data reveals more prominent high-value zones (marked by white circles in Fig. 9), indicating more defined spatial distribution characteristics. Furthermore, their frequency density distributions encompass more continuous and refined channel structures.

The realizations generated from 100 sets of randomly selected conditioning data exhibit a significant linear correlation with the corresponding references in terms of facies proportions (Fig. 10(a)). This indicates a high degree of consistency in facies distributions between the realizations and the references. Furthermore, in the two-dimensional MS-SWD-MDS scatter plots

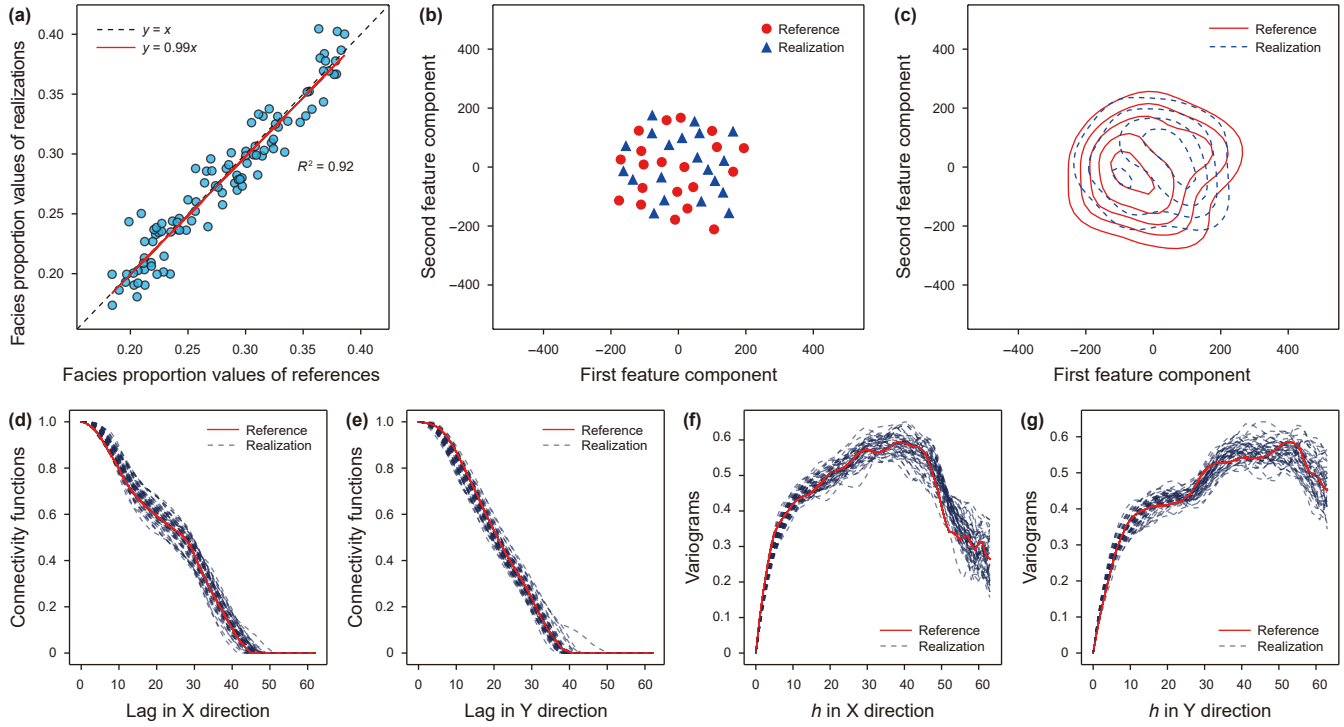

Fig. 10. Statistical comparison between realizations and references for the Brahmaputra River. **(a)** Facies proportions; **(b–c)** MS-SWD-MDS scatter plots and density plots; **(d–e)** connectivity functions in both X and Y directions; **(f–g)** variograms in both X and Y directions.

and density plots, the realizations and references demonstrate remarkable spatial correspondence, with their data distribution space showing substantial overlap (Fig. 10(b) and (c)). We obtained 50 realizations by simulating with fixed conditioning data. The results indicate that, in both the X and Y directions, the variograms and connectivity functions of the generated realizations closely mirror those of the reference (Fig. 10(d)–(g)), implying their comparable statistical properties. Specifically, these simulation results reaffirm that the proposed method can effectively integrate conditioning data, thereby significantly reducing the uncertainty inherent in stochastic simulations and enhancing the precision of sedimentary facies modeling.

4. Discussion

4.1. Interpretability of the framework

To elucidate the feature learning and decision-making mechanisms of the proposed method for sedimentary facies models, particularly the critical roles and contributions of each component, this section employs internal feature map analysis to dissect their dynamic interactions and attention patterns of these components across spatial extraction, transformation, and fusion, thereby enhancing the interpretability of the network model.

4.1.1. Importance of dual-domain analysis

This section systematically investigates the impact of different downsampling strategies on geological data representations by dissecting the DDLB. Comparative downsampling experiments involving Maxpool, Avgpool, and the Haar wavelet transform were designed to highlight the pivotal role of dual-domain feature analysis in sedimentary facies modeling. As a core component of the encoding path, the downsampling module learns conditioning data and captures essential features. As depicted in Fig. 11, the sampling process spans three scales, progressively downscaling with increasing network depth. At Scale 1, the Haar wavelet transform demonstrates superior feature extraction (Fig. 11(a) and (b)), accurately capturing the distribution of soft probability data (marked by black lines in Fig. 11(b)) and fully delineating channel facies wells (marked by white circles in Fig. 11(b)). Likewise, Maxpool effectively captures channel facies wells (marked by white circles in Fig. 11(c)) with wide high-value regions, demonstrating robust local feature extraction. However, as a local saliency selector, Maxpool exhibits limitations in extracting probability data features (Fig. 11(c)), with the right-side channel region consistently showing low-value characteristics (marked by the red circle in Fig. 11(c)). In contrast, Avgpool at this scale exhibits feature continuity (Fig. 11(d)), producing high-value regions with strong internal consistency and smooth transitions (marked

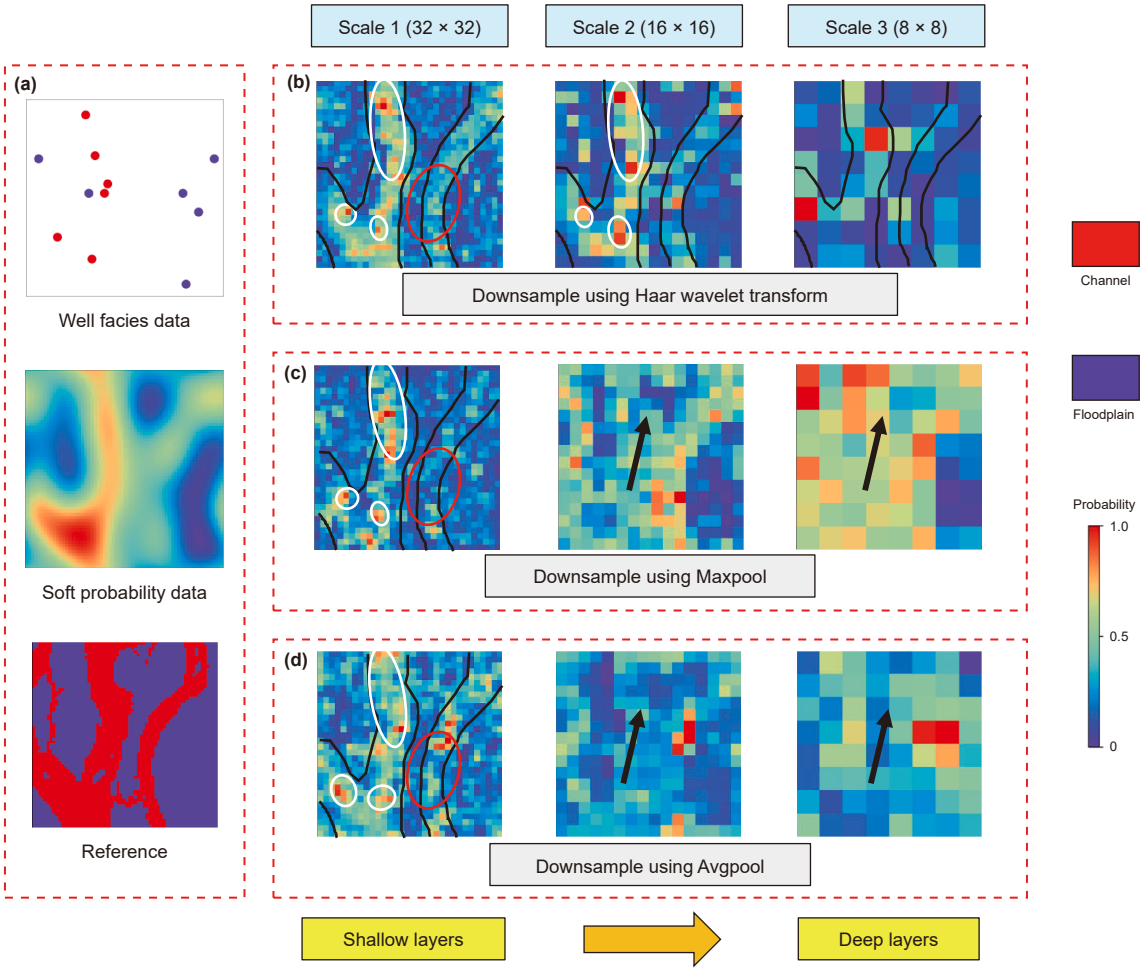

Fig. 11. Several downsampling approaches in DDLB. The Haar wavelet transform excels at extracting overall channel features with clear boundaries. In contrast, Maxpool and Avgpool tend to capture trend features.

by black lines in Fig. 11(d)), indicating its strength in capturing the overall distribution of geological data, though some localized low-value regions remain (marked by the red circle in Fig. 11(d)). As network depth increases (i.e., from Scale 2 to Scale 3), features extracted by the Haar wavelet transform maintain high consistency with both soft probability and well facies data (Fig. 11(b)). Traditional downsampling approaches (e.g., Maxpool and Avgpool) yield features with high-value regions but capture only local channel segments (Fig. 11(c) and (d)), with distribution trends roughly reflecting the general sedimentary facies patterns (marked by black arrows).

Experimental results demonstrate that the Haar wavelet transform, through layer-by-layer learning, accurately extracts global distribution patterns and certain refined geological details from conditioning data. The transform preserves key sedimentary patterns while effectively filtering noise. However, traditional downsampling approaches perform well in the spatial domain, capturing local geological features with more concentrated high-value regions (e.g., certain well facies data and soft probability data). By integrating the strengths of spatial domain and frequency domain feature learning, multi-dimensional synergy between local details and global patterns can be achieved. Figs. S4 and S5 in the Supporting material further highlight the necessity of multi-domain analysis. The results demonstrate that the multi-domain structure plays a critical role in improving simulation accuracy, as well as enhancing correlation and statistical significance (p-value) with respect to the reference model.

4.1.2. Performance of attention guide

(1) RSAG analysis

In this module, the RSAG serves as a feature concatenation bridge between the encoder and decoder, guiding the feature layout of the encoder using the high-level semantic features from the decoder. $RSAG_1$, $RSAG_2$, and $RSAG_3$ are embedded in the shallow, middle, and deep layers of the network, respectively. In the shallow $RSAG_1$, the attention features exhibit a discrete distribution, primarily focusing on channel facies wells and high-value regions of probability data (Fig. 12). As the network deepens to the $RSAG_2$, high-value regions form a continuous distribution, initially revealing the spatial patterns (Fig. 12). Finally, in the deep-layer $RSAG_3$, the feature map displays clear boundaries and continuous high-value regions, accurately characterizing the spatial distribution of sedimentary patterns consistent with the reference.

(2) PMSA analysis

The PMSA module, positioned at the end of the decoding path, achieves high-precision geological model reconstruction through multi-scale feature fusion (Fig. 13). The decoder extracts low-resolution features (F_1 and F_2) with high abstraction, revealing isolated patchy structures, while high-resolution features (F_3 and F_4) are sparsely distributed, with their corresponding sedimentary facies models predominantly developing in medium-to-high-value regions (details of F_1–F_4 are shown in Fig. 13). After multi-scale fusion, isolated channels exhibit improved connectivity, indicating that integrating structural features from multiple scales can compensate for the deficiencies of single-scale information. Owing to its unique cross-spatial feature extraction mechanism, PMSA demonstrates superior performance by effectively capturing both long-range and short-range spatial dependencies. This capability allows for the precise identification of key sedimentary patterns and the generation of attention maps with enhanced spatial continuity and sharper boundary delineation. This effectively suppresses background noise while significantly enhancing simulation accuracy and reliability. The optimized geological features exhibit strong spatial correspondence with the generated realization, confirming the capability of PMSA to transform abstract, discrete multi-scale features into continuous, realistic geological representations.

4.2. Pattern-driven performance and simulation stability

In sedimentary facies modeling, the core of the proposed method lies in fitting complex patterns rather than directly matching conditioning data. Specifically, the generated

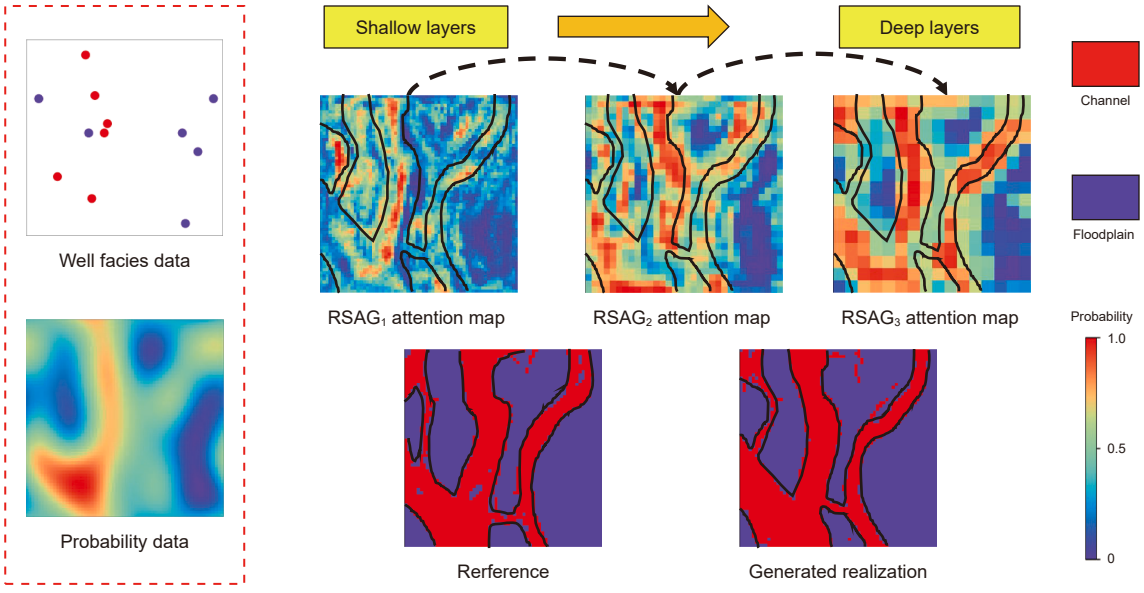

Fig. 12. Feature learning of RSAG. From the attention feature maps at different stages, it is evident that in shallow layers, RSAG primarily focuses on the preliminary learning of conditioning data distributions, with its attention mechanism prominently concentrated on high-value regions of channel facies wells and probability data.

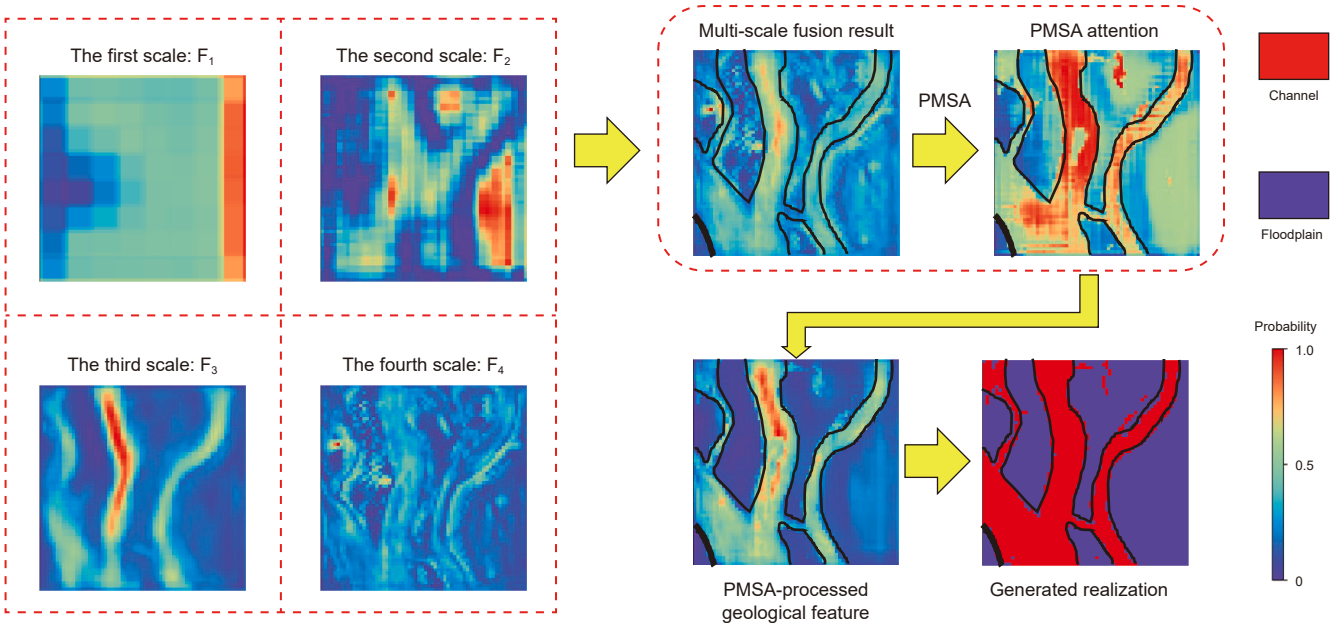

Fig. 13. Feature learning of PMSA. It elucidates the critical role of the PMSA module in multi-scale fusion learning: the multi-scale fusion results exhibit contiguous distributions of medium-to-high-value regions, while the high-value regions in the PMSA attention feature maps precisely delineate the spatial distribution characteristics of channels.

realizations must satisfy multiple constraints: they must accurately align with well data while preserving reasonable geological characteristics. Furthermore, the simulation results should be broadly consistent with the distribution features of soft probability data. This pattern-driven method offers dual advantages: it mitigates model collapse due to overfitting conditioning data and ensures the reliability and generalizability of the sedimentary facies models. To validate the performance of the proposed method, we conducted a series of comparative experiments using the two datasets presented in Fig. 6.

Given that well facies data are typically spatially sparse and soft probability data exhibit uneven quality and high variability, this study focuses on evaluating the performance of sedimentary facies modeling under varying soft probability data quality constraints. Fig. 14 illustrates three cases for two datasets, with soft probability data of different qualities generated by setting reconstruction coefficient R to 2, 2.5, 4, and 5. For all cases at $R = 2$ in Fig. 14, the soft probability data quality is poor, making it challenging to establish a direct, reliable correspondence with the references. Notably, in Fig. 14(b), the soft probability data for Case 1 and Case 3 at $R = 2$ are highly similar (marked by black boxes), yet their corresponding references exhibit significant differences. This phenomenon clearly highlights the substantial uncertainty in mapping low-quality probability data to sedimentary facies models. In the columns corresponding to $R = 2$, the generated results for the conceptual river and Brahmaputra River datasets show slight deviations in detailed distributions compared to the references, but their overall pattern characteristics remain similar (marked by the yellow boxes). This outcome arises because low-quality soft probability data ($R = 2$) fail to adequately capture the topological characteristics of the references. Although some deviations exist in finer details, the simulated realizations still exhibit good consistency with the references in terms of channel width and sinuosity. The quality of conditioning data undoubtedly influences realizations to a certain extent. Lower-resolution soft probability data inevitably increase result deviations, which is the inherent limitation. However, the performance of the proposed method does not solely depend on the quality of the conditioning data. Its core strength lies in two key aspects: (1) effectively extracting useful feature information from the available conditioning data, and (2) accurately reproducing complex geological patterns. Therefore, when the resolution of conditioning data is low and extractable information is limited, the method's capability in pattern fitting and reproduction becomes particularly crucial. Owing to its outstanding pattern reproduction capability, the proposed method still produces realizations that conform well to the overall geological pattern under low-resolution constraints and exhibits no pattern collapse. This demonstrates that the proposed method possesses excellent generalization capability and robustness.

As the value of R increases progressively to 2.5, 4, and 5, the similarity between the generated models and the references improves significantly. Notably, in the columns corresponding to $R = 5$, the simulation results for both datasets exhibit high consistency with the references in terms of both overall morphology and local features (Fig. 14(a) and (b)), including channel sinuosity, branching and confluence patterns, and the scale and distribution of fine channel structures (highlighted by the white boxes). This indicates that the method not only effectively extracts information from the available conditioning data, but also possesses a strong capability to accurately reproduce complex spatial geological patterns. The Brahmaputra River dataset is derived from binarized satellite remote sensing imagery (Fig. S3 in the Supporting material). Its geometric characteristics, including channel bifurcation and confluence patterns, multi-scale channel width distributions, and sinuosity, are highly consistent with the sedimentary dynamics of modern fluvial systems. Under conditioning constraints of varying soft probability data quality, the generated realizations maintain strong agreement with the references in terms of overall non-stationary pattern characteristics (e.g., variations in channel width and sinuosity). This demonstrates that the proposed method can effectively reproduce sedimentary facies that conform to established principles of sedimentary dynamics. To further evaluate the model's generalization ability, we conducted additional noise robustness experiments. The results show that, even under strong Gaussian filtering interference, the generated

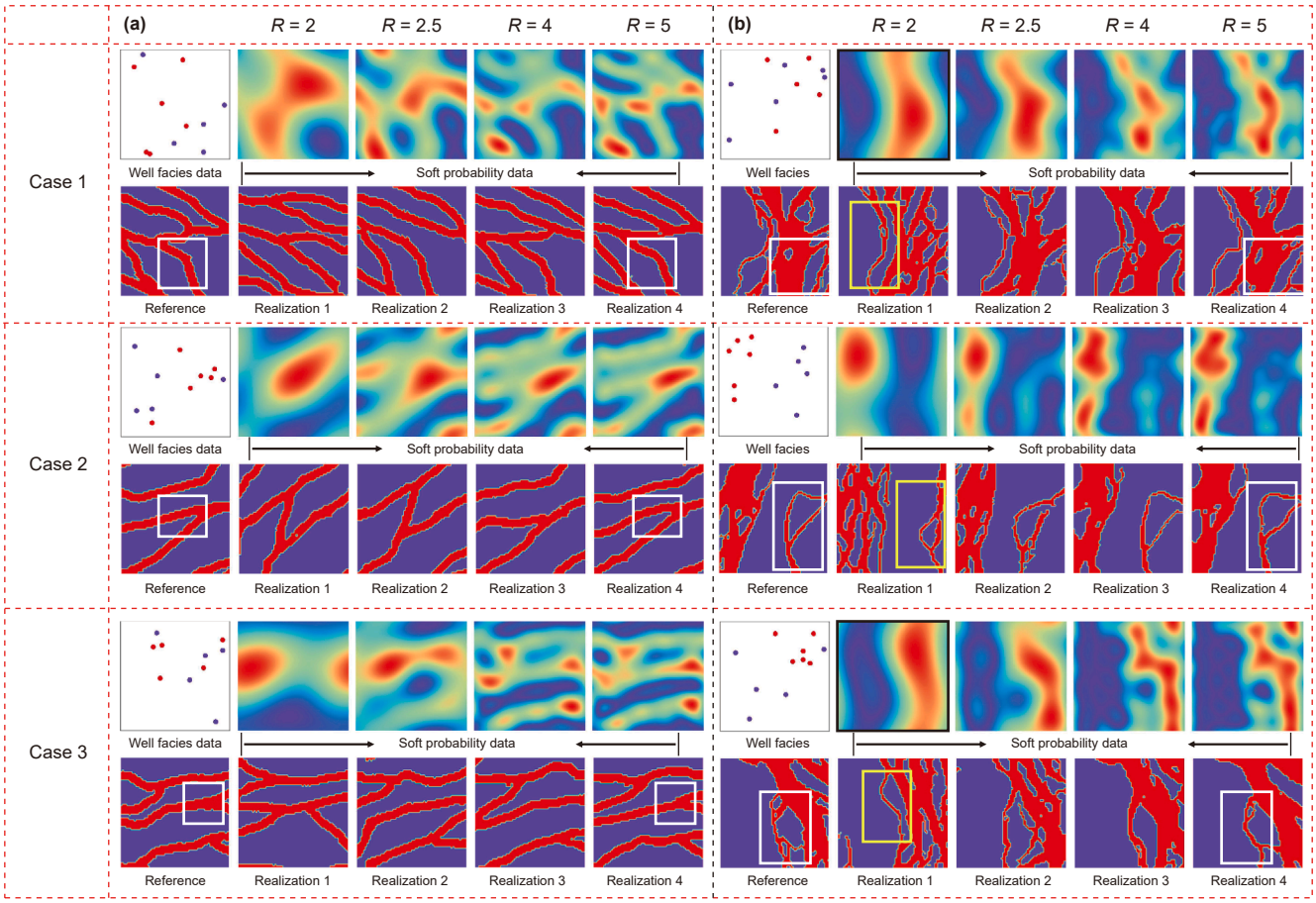

Fig. 14. Simulation cases for the datasets generated under well facies data and varying soft probability data constraints. **(a)** Conceptual river; **(b)** the Brahmaputra River. Soft probability data at $R = 2$ represents the lowest-quality probability constraints, while that at $R = 5$ denotes the highest-quality constraints.

realizations still strictly adhere to the distribution characteristics of the river channels (Fig. S6 in the Supporting material), fully demonstrating the model's robust pattern constraint capability and generalization performance.

To quantitatively evaluate the performance of the proposed method in reproducing sedimentary facies models, the Brahmaputra River dataset was analyzed, which has complex pattern characteristics. Facies proportion distributions of 50 realizations generated at $R = 2$, 2.5, 4, and 5 were statistically compared (Fig. 15(a)). As R decreases (i.e., the quality of soft probability data gradually decreases), the range of corresponding facies proportions tends to widen. Nevertheless, the overall distribution of facies proportions in the generated realizations at $R = 2$ and 2.5 still shows high consistency with that of the references. This phenomenon demonstrates that the method can effectively learn and adhere to the intrinsic governing rules of geological patterns, thereby serving as a critical constraint and guiding mechanism during the modeling process. This mechanism significantly contributes to reducing uncertainty in the realizations. These results confirm the method's robust adaptability to varying soft probability data quality constraints and its superior generalization ability in reproducing pattern characteristics.

Furthermore, 2D MS-SWD-MDS plots were employed to visually illustrate the distribution of different data in the feature space (Fig. 15(b) and (c)). To explicitly compare the impact of soft probability data quality on the feature space of generated realizations, the density distributions at $R = 2$ and 5 were mapped from Fig. 15(c) onto the data space in Fig. 15(b). The results reveal that the distribution range for $R = 2$ almost entirely encompasses the references marked by red scatter points (Fig. 15(b)), indicating that low-quality soft probability data has a limited impact on the data representation space of the proposed method. At $R = 5$, the generated realizations not only fully encompass the red reference scatter points in feature space, but also exhibit dense clustering along the boundary of their distribution, indicating excellent spatial consistency with the reference patterns. These experimental results and analyses robustly demonstrate that the proposed method efficiently generates critical features based on limited well data and soft probability data. Notably, the sedimentary facies models generated by the method do not merely rely on direct constraints from conditioning data but leverage deep neural networks to uncover intrinsic data patterns for effective pattern fitting. Even under low-resolution soft probability data constraints, the model consistently generates simulations that align with the distributional trends of conditioning data while accurately reproducing the spatial distribution pattern of sedimentary facies models, thereby significantly reducing uncertainty in sedimentary facies models.

4.3. Advantages of the proposed method

Through feature map analysis in Section 4.1, the simulation process was dissected, clearly explaining the role of each component. In this section, the effectiveness of each component will be further validated through ablation experiments and comparative analysis. The comparative analysis methods used in this study can

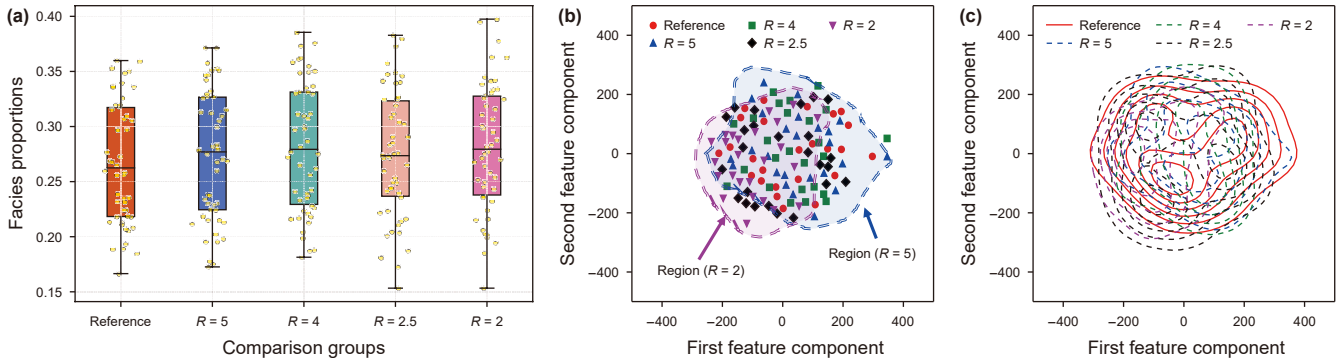

Fig. 15. Quantitative analysis results for realizations generated from the Brahmaputra River dataset under varying soft probability data constraints. **(a)** Plots of facies proportions; **(b–c)** MS-SWD-MDS scatter plots and density plots.

be primarily divided into two categories. The first involves an ablation study, systematically retaining or removing the attention-guided structures (i.e., RSAG and PMSA) and DDLBs to assess their individual contributions. The second category utilizes Direct Sampling, a classical multiple-point geostatistics technique widely adopted in sedimentary facies modeling (Mariethoz et al., 2010; Rezaee et al., 2013). Notably, DA-Net serves as the primary network model in this study, with all other methods configured as control groups for comparison. All methods incorporate joint constraints on well facies data and soft probability data during the simulation. To thoroughly evaluate the performance of the proposed model in extracting and reproducing complex patterns, this study conducts a comparative experimental analysis using the Brahmaputra River dataset, which is characterized by significant heterogeneity.

In the experiment, all network models were uniformly trained for 50 epochs, with parameter settings consistent with those in Section 3.2. A total of 30 independent test samples were randomly sampled, and corresponding conditioning data were generated for each sample. Based on these, 30 sets of simulation realizations were produced using DA-Net, AS-Net, DDLB-Net, Basic-Net, and Direct Sampling, with four representative cases illustrated in Fig. 16.

Compared to the traditional Direct Sampling method, the sedimentary facies models generated by the deep learning methods demonstrate significant advantages in morphological characteristics and spatial continuity (Fig. 16). This finding highlights the limitations of multiple-point geostatistics in extracting complex patterns and reconstructing fine-scale structures, while convolutional neural network architectures exhibit superior capabilities in pattern capture and representations. While Basic-Net can reproduce relatively clear channel facies, its basic network architecture and unoptimized feature learning mechanisms limit its accuracy in reproducing features such as channel distribution patterns. In contrast, AS-Net and DDLB-Net build upon Basic-Net by incorporating the attention-guided structures and DDLBs, respectively, thereby enhancing pattern feature representations. DDLB-Net employs the dual-domain learning strategy, jointly extracting spatial and frequency domain features to effectively capture both local and global conditions, thereby reducing

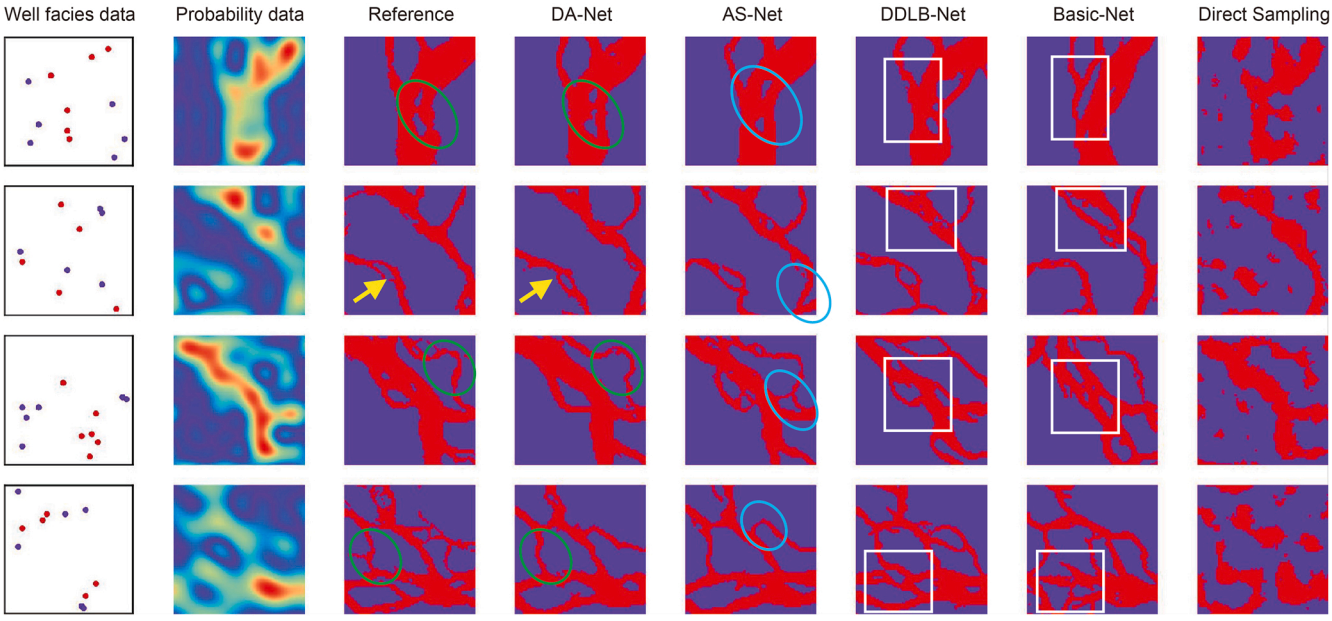

Fig. 16. Comparison of geological realizations generated by different methods based on the Brahmaputra River dataset. Compared to Direct Sampling method, the deep learning methods (e.g., DA-Net, AS-Net, DDLB-Net, and Basic-Net) produce realizations with higher consistency in pattern characteristics and morphological structures relative to the references.

uncertainty in simulation results. Compared to Basic-Net, DDLB-Net generates sedimentary facies models with improved spatial distribution patterns and local features (marked in white boxes). AS-Net, leveraging an attention mechanism, uses features recovered by the decoder to guide the encoder in extracting conditional features. Combined with multi-scale cross-spatial fusion, this enables the capture of long-range dependencies. Comprehensive comparative analysis reveals that DA-Net, by integrating attention-guided structures and dual-domain learning strategies, achieves the best consistency with the references in both local details and overall distribution features (marked by green circles). Notably, due to data threshold segmentation, the reference exhibits discontinuities at certain fine channel connections (indicated by the yellow arrows), and DA-Net intelligently infers and reconstructs consistent sedimentary facies models by learning sedimentary pattern features (Fig. 16), thereby fully demonstrating its superiority in extracting and reproducing complex spatial distribution patterns.

To systematically evaluate model performance, this study employs multi-point density functions for quantitative analysis of various models (Fig. 17(a)). The results indicate that the realizations generated by DA-Net exhibit high consistency with the references across 16 statistical templates. Fig. 17(b) further compares the spatial variability characteristics of different models at template 0 and 15. Comparative analysis reveals that, compared to Basic-Net, the incorporation of attention-guided structures and the DDLBs significantly improves the alignment of simulated spatial variability features with the references. Among these evaluation indicators, DA-Net performs the best, demonstrating better statistical consistency with the references across different templates. Notably, Basic-Net still outperforms the traditional Direct Sampling in multi-point pattern reproduction. To clearly assess the performance improvement of the proposed method over the traditional geostatistical method, this study focuses on comparing DA-Net with Direct Sampling, particularly emphasizing their ability to capture the data distribution characteristics of realizations. As shown in Fig. S7 of the Supporting material, the realizations produced by DA-Net exhibit a significantly higher degree of distributional overlap with the references. In contrast, the realizations generated by Direct Sampling show only partial alignment with the references, indicating that this method can only capture and reproduce a subset of the structural features of sedimentary facies. These results robustly confirm the pronounced advantages of our proposed method in mining complex sedimentary patterns and reconstructing accurate realizations. Additionally, we introduced similarity evaluation indicators, including Wasserstein distance, KID and K-MMD, to further refine the comparison of pattern reproduction performance and simulation accuracy across all methods. As shown in Table 1, deep learning models generally perform well, with DA-Net achieving the best results across Wasserstein distance (0.09), KID (0.0017) and K-MMD (0.21), indicating highly similar distributions between the generated geological realizations and the references. In contrast, Direct Sampling performs the poorest across these indicators (Wasserstein distance: 0.9, KID: 0.1738, K-MMD: 1.03). This is attributed to its lack of complex sedimentary pattern representation capabilities and reliance solely on the direct statistical method. Comprehensive analysis confirms that deep learning models offer significant advantages in sedimentary facies models, with DA-Net particularly excelling in the characterization of complex sedimentary facies models, providing a reliable technical solution for fine-scale sedimentary facies modeling. Furthermore, the computational cost statistics (Table S5 in the Supporting material) indicate that, despite the incorporation of DDLB and attention-guided structures, our method does not result in significant increases in training time, GPU memory usage, or inference time, while achieving substantial improvements in simulation performance.

4.4. Application prospects and limitations and future work

This study proposes an interpretable GAN framework that integrates attention-guided mechanisms with dual-domain learning to efficiently capture heterogeneous sedimentary patterns. The proposed method demonstrates multifaceted technical advantages in practical geological applications. Once trained, the proposed method achieves efficient conditional modeling, greatly accelerating the modeling process. In hydrocarbon exploration, it generates high-fidelity reservoir models, providing a reliable basis for reservoir prediction and optimization of development strategies (e.g., well placement optimization). In the field of carbon sequestration, the precise reconstruction of sedimentary facies distributions facilitates the identification of suitable storage sites. For groundwater evaluation, the sedimentary facies realizations

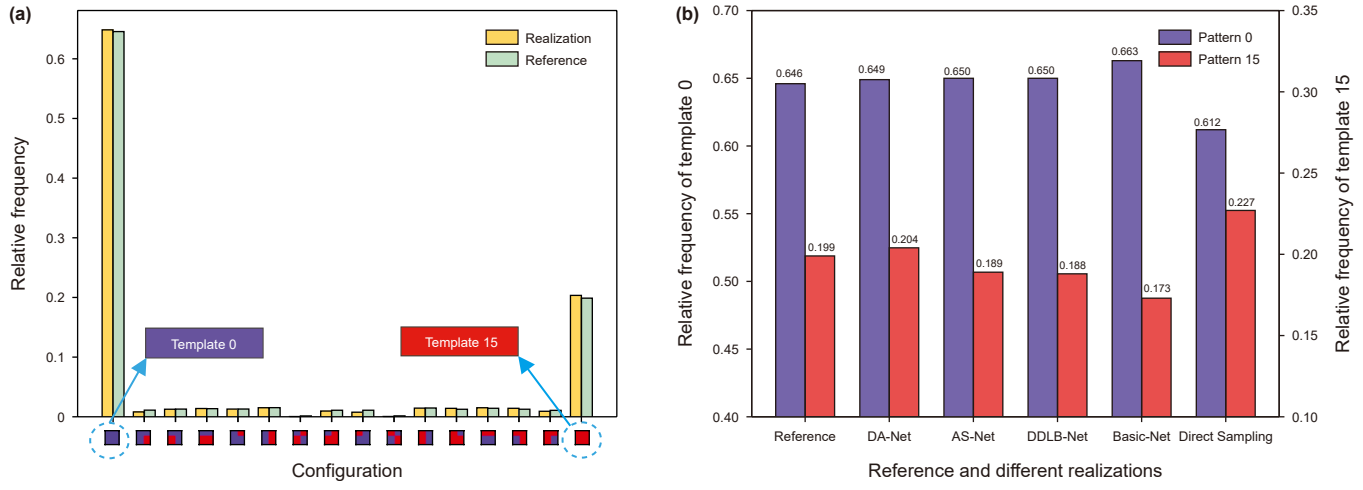

Fig. 17. Statistical results from multiple-point density functions. **(a)** Comparison of multi-point pattern statistical features between realizations generated by the DA-Net and the references; **(b)** multi-point pattern statistical features of different simulation methods at template 0 and template 15.

Table 1
The quality evaluation of the experimental results for each algorithm, where smaller values indicate greater similarity to the references. The best results (shown in bold) are achieved by DA-Net. In the table, × indicates that the block is not included, √ indicates that the block is included, and ↓ means that lower values correspond to more realistic realizations. / indicates that the method has no corresponding structures.

Backbone	DDLB	Attention-guided structures	Name	Wasserstein distance (↓)	KID (↓)	K-MMD (↓)
Basic-Net	×	×	Basic-Net	0.58	0.0124	0.30
	√	×	DDLB-Net	0.16	0.0087	0.27
	×	√	AS-Net	0.23	0.0032	0.25
	√	√	DA-Net	**0.09**	**0.0017**	**0.21**
/	/	/	Direct Sampling	0.90	0.1738	1.03

generated by the method support groundwater flow simulations and resource management. Furthermore, the interpretability of the method enables intuitive identification of the modeling process. This not only deepens the understanding of the model's decision-making mechanisms but also strengthens the credibility of the simulation results in geological applications. This method demonstrates strong practicality. Its soft probability constraints can effectively accommodate typical geophysical data, enabling the rapid establishment of efficient workflows (Fig. S8 in the Supporting material). The proposed framework integrates multi-source conditional constraints with deep generative modeling. Its core design is independent of specific depositional systems, thereby exhibiting excellent transferability and broad applicability.

Despite the suboptimal quality of soft probability data, the DA-Net remains capable of reconstructing reasonable geological realizations (Fig. 14). However, the simulation accuracy of the network model is still constrained by the quality of conditioning data, particularly in capturing details of sedimentary facies models, where the results exhibit inherent probabilistic randomness and uncertainty. This is primarily attributed to the absence of strict geological constraints. To address this limitation, integrating multi-source heterogeneous data (e.g., production dynamics time-series data and high-resolution core data) and prior geological knowledge (e.g., provenance characteristics and empirical formulas), and encoding these as differentiable constraint conditions, offers an effective research paradigm to overcome such challenges. Specifically, the essence of the research paradigm lies in incorporating additional prior conditions and knowledge to further constrain the possibility space of simulation realizations (Hu et al., 2024; Song et al., 2023). In the future, we will advance this research paradigm, focusing on achieving an organic integration and synergistic optimization of cross-modal representation learning and knowledge inference to further enhance the precision of geological characterization.

5. Conclusions

Focusing on the critical issue of refined sedimentary facies model characterization, this study proposes an interpretable attention-guided GAN framework with dual-domain learning. Through extensive experiments, comprehensive evaluation, and validation, the framework demonstrates superior simulation performance. The generated sedimentary facies models exhibit high geological fidelity, accurately reproducing key geological characteristics, while also showcasing realistic spatial heterogeneity and finely reconstructing intricate details of complex sedimentary facies models. These advancements are driven by the following key aspects:

(1) Data mining across multiple domains is necessary. Through a dual-domain collaborative processing mechanism, the framework can deeply explore the prior information of sedimentary facies models from multiple perspectives and effectively analyze complex geological features, thereby significantly improving the accuracy of sedimentary facies modeling.
(2) Cross-spatial attention mechanisms guide the model in capturing dependencies and critical patterns of sedimentary facies models. This guidance is reflected in two main aspects: (i) effectively integrating multi-resolution information within cascaded encoders and decoders; and (ii) focusing on the cross-spatial geological features at different scales, embedding spatial positional information to capture both short-range and long-range dependency patterns.
(3) The hybrid loss functions integrate gradient penalty, hinge loss, and conditional loss through weighted coupling to formulate a synergistic optimization task. Simulation experiments demonstrate that the dynamic weighted coupling of these loss components enables the generated geological realizations to achieve both high geological realism and strong adherence to geological constraints.
(4) The proposed interpretable framework is valuable, and it has shown success in enhancing the interpretability of deep learning by providing clear insights into the complete process of conditioning data encoding and geological model decoding. Additionally, this interpretable study offers valuable insights into the practical operations of each module.

CRediT authorship contribution statement

Lei Liu: Writing – review & editing, Writing – original draft, Software, Methodology. **Wei Li:** Writing – review & editing, Validation. **Jian Gao:** Supervision, Project administration. **Da-Li Yue:** Writing – review & editing, Project administration. **De-Gang Wu:** Validation, Methodology. **Wu-Rong Wang:** Validation, Supervision. **Jin Lin:** Visualization, Validation. **Zhi-Bo Li:** Validation, Software. **Qian Zhong:** Visualization, Methodology. **Jia-Gen Hou:** Supervision.

Declaration of competing interest

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

Acknowledgements

This work was supported by National Science and Technology Major Project "CO_2 Flooding for Significantly Enhancing Recovery Rate and Long-Term Sequestration Technology" (No. 2024ZD1406601), National Natural Science Foundation of China (Nos. 42272186, 42472179, 42302128, 42202109), Frontier Interdisciplinary Exploration Research Program of China University of Petroleum, Beijing (No. 2462024XKQY003), and Science Foundation

of China University of Petroleum (Beijing) (Nos. 2462023BJRC024, and 2462023YJRC039).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.petsci.2026.02.018.

References

Abdellatif, A., Elsheikh, A.H., Busby, D., Berthet, P., 2023. Generation of non-stationary stochastic fields using Generative Adversarial Networks. arXiv arXiv: 2205.05469v2 [cs.LG]. https://doi.org/10.48550/arXiv.2205.05469.

Amidror, I., 2013. Mastering the Discrete Fourier Transform in One, Two or Several Dimensions: Pitfalls and Artifacts, Computational Imaging and Vision. Springer, London, London. https://doi.org/10.1007/978-1-4471-5167-8.

Chen, W., Zhong, C., Qin, X., Wang, L., 2023. Intelligent Interpretation for Geological Disasters: From space-air-ground Integration Perspective. Springer Nature, Singapore, Singapore. https://doi.org/10.1007/978-981-99-5822-1.

Chi, P., Sun, J., Zhang, R., Yan, W., Dong, H., Cui, L., Cui, R., Luo, X., 2025. Multidimensional data-driven porous media reconstruction: Inversion from 1D/2D pore parameters to 3D real pores. Pet. Sci. https://doi.org/10.1016/j.petsci.2025.03.046.

Cui, Z., Chen, Q., Liu, G., 2023. A two-stage downscaling hydrological modeling approach via convolutional conditional neural process and geostatistical bias correction. J. Hydrol. 620, 129498. https://doi.org/10.1016/j.jhydrol.2023.129498.

Cui, Z., Chen, Q., Liu, G., Xun, L., 2024a. SA-RelayGANs: a novel framework for the characterization of complex hydrological structures based on GANs and self-attention mechanism. Water Resour. Res. 60, e2023WR035932. https://doi.org/10.1029/2023WR035932.

Cui, Z., Chen, Q., Luo, J., Ma, X., Liu, G., 2024b. Characterizing subsurface structures from hard and soft data with multiple-condition fusion neural network. Water Resour. Res. 60, e2024WR038170. https://doi.org/10.1029/2024WR038170.

Fan, W., Liu, G., Chen, Q., Cui, Z., Fang, H., Chen, G., Wu, X., 2024a. Automatic reconstruction of geological reservoir models based on conditioning data constraints and BicycleGAN. Geoenergy Sci. Eng. 234, 212690. https://doi.org/10.1016/j.geoen.2024.212690.

Fan, W., Liu, G., Chen, Q., Cui, Z., Wu, X., Zhang, Z., 2024b. Stochastic reconstruction of geological reservoir models based on a concurrent multi-stage U-Net generative adversarial network. Comput. Geosci. 186, 105562. https://doi.org/10.1016/j.cageo.2024.105562.

Feng, R., Grana, D., Mosegaard, K., 2025. Geostatistical facies simulation based on training image using generative networks and gradual deformation. Math. Geosci. 57, 1021–1044. https://doi.org/10.1007/s11004-024-10169-y.

Feng, R., Grana, D., Mukerji, T., Mosegaard, K., 2022. Application of bayesian generative adversarial networks to geological facies modeling. Math. Geosci. 54, 831–855. https://doi.org/10.1007/s11004-022-09994-w.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C., 2017. Improved training of Wasserstein GANs. arXiv arXiv:1704.00028v3 [cs.LG]. https://doi.org/10.48550/arXiv.1704.00028.

Hu, X., Song, S., Hou, J., Yin, Y., Hou, M., Azevedo, L., 2024. Stochastic modeling of thin mud drapes inside point bar reservoirs with ALLUVSIM-GANSim. Water Resour. Res. 60 e2023WR035989. https://doi.org/10.1029/2023WR035989.

Jona Lasinio, G., Pyrcz, Michael J., Deutsch, Clayton V., 2015. Geostatistical reservoir modeling. Metron 73, 149–150. https://doi.org/10.1007/s40300-014-0058-8.

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, 381–406. https://doi.org/10.1002/2017WR022148.

Li, X., Mariethoz, G., Lu, D., Linde, N., 2016. Patch-based iterative conditional geostatistical simulation using graph cuts. Water Resour. Res. 52, 6297–6320. https://doi.org/10.1002/2015WR018378.

Liu, J., Liu, J., 2022. Integrating deep learning and logging data analytics for lithofacies classification and 3D modeling of tight sandstone reservoirs. Geosci. Front. 13, 101311. https://doi.org/10.1016/j.gsf.2021.101311.

Liu, L., Yue, D., Li, W., Wu, D., Gao, J., Zhong, Q., Wang, W., Hou, J., 2025. A novel stochastic simulation method for sedimentary facies based on the generative adversarial network with a spatially-adaptive conditioning module and comprehensive attention mechanisms. Geoenergy Sci. Eng. 249, 213758. https://doi.org/10.1016/j.geoen.2025.213758.

Liu, X., Chen, X., Cheng, J., Zhou, L., Chen, L., Li, C., Zu, S., 2023. Simulation of complex geological architectures based on multistage generative adversarial networks integrating with attention mechanism and spectral normalization. IEEE Trans. Geosci. Rem. Sens. 61, 1–15. https://doi.org/10.1109/TGRS.2023.3294493.

Liu, X., Cheng, J., Cai, Y., Mo, Q., Li, C., Zu, S., 2022. Stochastic simulation of facies using deep convolutional generative adversarial network and image quilting. Mar. Petrol. Geol. 146, 105932. https://doi.org/10.1016/j.marpetgeo.2022.105932.

Mariethoz, G., Renard, P., Straubhaar, J., 2010. The Direct Sampling method to perform multiple-point geostatistical simulations. Water Resour. Res. 46, 2008WR007621. https://doi.org/10.1029/2008WR007621.

Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D., 2018. Attention U-Net: learning where to look for the pancreas. arXiv arXiv: 1804.03999. https://doi.org/10.48550/arXiv.1804.03999.

Ouyang, D., He, S., Zhang, G., Luo, M., Guo, H., Zhan, J., Huang, Z., 2023. In: Efficient multi-scale attention module with cross-spatial learning. ICASSP, Rhodes Island, Greece, pp. 1–5.

Pan, W., Torres-Verdín, C., Pyrcz, M.J., 2021. Stochastic pix2pix: A new machine learning method for geophysical and well conditioning of rule-based channel reservoir models. Nat. Resour. Res. 30, 1319–1345. https://doi.org/10.1007/s11053-020-09778-1.

Pawar, R.J., Bromhal, G.S., Chu, S., Dilmore, R.M., Oldenburg, C.M., Stauffer, P.H., Zhang, Y., Guthrie, G.D., 2016. The national risk assessment partnership's integrated assessment model for carbon storage: A tool to support decision making amidst uncertainty. Int. J. Greenh. Gas Control 52, 175–189. https://doi.org/10.1016/j.ijggc.2016.06.015.

Rezaee, H., Mariethoz, G., Koneshloo, M., Asghari, O., 2013. Multiple-point geostatistical simulation using the bunch-pasting direct sampling method. Comput. Geosci. 54, 293–308. https://doi.org/10.1016/j.cageo.2013.01.020.

Ruffino, C., Hérault, R., Laloy, E., Gasso, G., 2020. Pixel-wise conditioned generative adversarial networks for image synthesis and completion. Neurocomputing 416, 218–230. https://doi.org/10.1016/j.neucom.2019.11.116.

Song, S., Mukerji, T., Hou, J., 2022a. Bridging the gap between geophysics and geology with generative adversarial networks. IEEE Trans. Geosci. Rem. Sens. 60, 1–11. https://doi.org/10.1109/TGRS.2021.3066975.

Song, S., Mukerji, T., Hou, J., 2021a. Geological facies modeling based on progressive growing of generative adversarial networks (GANs). Comput. Geosci. 25, 1251–1273. https://doi.org/10.1007/s10596-021-10059-w.

Song, S., Mukerji, T., Hou, J., 2021b. GANSim: conditional facies simulation using an improved progressive growing of generative adversarial networks (GANs). Math. Geosci. 53, 1413–1444. https://doi.org/10.1007/s11004-021-09934-0.

Song, S., Mukerji, T., Hou, J., Zhang, D., Lyu, X., 2022b. GANSim-3D for conditional geomodelling: theory and field application. Water Resour. Res. 58, e2021WR031865. https://doi.org/10.1029/2021WR031865.

Song, S., Zhang, D., Mukerji, T., Wang, N., 2023. GANSim-surrogate: An integrated framework for stochastic conditional geomodelling. J. Hydrol. 620, 129493. https://doi.org/10.1016/j.jhydrol.2023.129493.

Sun, C., Demyanov, V., Arnold, D., 2023. A conditional GAN-based approach to build 3D facies models sequentially upwards. Comput. Geosci. 181, 105460. https://doi.org/10.1016/j.cageo.2023.105460.

Tetteh, M., Li, L., Davis, A., 2024. Leveraging deep learning with progressive growing GAN and ensemble smoother with multiple data assimilation for inverse modeling. Adv. Water Resour. 187, 104680. https://doi.org/10.1016/j.advwatres.2024.104680.

Wang, Y., Dai, Z., Wang, G., Chen, L., Xia, Y., Zhou, Y., 2024. A hybrid physics-informed data-driven neural network for CO2 storage in depleted shale reservoirs. Pet. Sci. 21, 286–301. https://doi.org/10.1016/j.petsci.2023.08.032.

Xu, G., Liao, W., Zhang, X., Li, C., He, X., Wu, X., 2023. Haar wavelet downsampling: A simple but effective downsampling module for semantic segmentation. Pattern Recogn. 143, 109819. https://doi.org/10.1016/j.patcog.2023.109819.

Zakeri, F., Mariethoz, G., 2021. A review of geostatistical simulation models applied to satellite remote sensing: Methods and applications. Remote Sens. Environ. 259, 112381. https://doi.org/10.1016/j.rse.2021.112381.

Zhan, C., Dai, Z., Samper, J., Yin, S., Ershadnia, R., Zhang, X., Wang, Y., Yang, Z., Luan, X., Soltanian, M.R., 2022. An integrated inversion framework for heterogeneous aquifer structure identification with single-sample generative adversarial network. J. Hydrol. 610, 127844. https://doi.org/10.1016/j.jhydrol.2022.127844.

Zhang, C., Song, X., Azevedo, L., 2021. U-net generative adversarial network for subsurface facies modeling. Comput. Geosci. 25, 553–573. https://doi.org/10.1007/s10596-020-10027-w.

Zhang, K., Yu, H., Ma, X., Zhang, J., Wang, J., Yao, C., Yang, Y., Sun, H., Yao, J., Wang, J., 2022. Multi-source information fused generative adversarial network model and data assimilation based history matching for reservoir with complex geologies. Pet. Sci. 19, 707–719. https://doi.org/10.1016/j.petsci.2021.10.007.

Zhang, T., Bian, N., Liu, Q., Du, Y., 2024. 3D super-resolution reconstruction of porous media based on GANs and CBAMs. Stoch. Environ. Res. Risk Assess. 38, 1475–1504. https://doi.org/10.1007/s00477-023-02639-2.

Zhang, T., Tilke, P., Dupont, E., Zhu, L., Liang, L., Bailey, W., 2019. Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks. Pet. Sci. 16, 541–549. https://doi.org/10.1007/s12182-019-0328-4.