



China Geology

Journal homepage: <http://chinageology.cgs.cn>
<https://www.sciencedirect.com/journal/china-geology>



Deep learning framework with hybrid loss function for seismic velocity inversion of natural gas hydrate reservoirs

Zhen-wei Guo^{a, b, c}, Jun Lei^{a, b, c}, Bo-chen Wang^{a, b, c, *}, Yan-yi Wang^{a, b, c}, Yan-jun Chen^{a, b, c}, Jian-xin Liu^{a, b, c}

^a School of Geosciences and Info-physics, Central South University, Changsha 410083, China

^b Key Laboratory of Non-Ferrous Resources and Geological Hazard Detection, Changsha 410083, China

^c Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring, Ministry of Education, Changsha 410083, China

ARTICLE INFO

Article history:

Received 29 September 2025

Received in revised form 27 February 2026

Accepted 3 April 2026

Available online 7 May 2026

Keywords:

Natural gas hydrate
 Clean energy
 Combustible ice
 Unconventional energy
 Seismic velocity inversion
 Deep learning
 Hybrid loss function
 U-Net-like architecture
 Full waveform inversion
 Gas hydrate exploration
 Risk assessment

ABSTRACT

Natural gas hydrate (NGH) has attracted increasing attention as a promising unconventional energy resource owing to its high volumetric storage capacity, yet its development is accompanied by significant greenhouse gas risks. Therefore, accurate reservoir characterization is vital for marine resource exploration and sustainable development. Full waveform inversion (FWI) offers high-resolution imaging, yet suffers from heavy computation, sensitivity to initial models and non-uniqueness. Recent deep learning (DL) methods improve efficiency and accuracy, nevertheless, still struggle with clear boundary extraction and multi-level semantic representation. The authors propose a novel deep architecture (SC-UNeXt) designed to learn a mapping from seismic records to velocity models, which integrates a U-Net backbone, ConvNeXt residual blocks, spatial-channel squeeze-and-excitation attention and pixel shuffle up-sampling. Furthermore, a hybrid loss integrating mean squared error, multi-scale structural similarity, and perceptual discrepancy simultaneously optimizes pixel-wise accuracy, structural fidelity, and semantic consistency. Comprehensive tests on both synthetic NGH data and the 3D SEG/EAGE marine overthrust model with NGH demonstrate that SC-UNeXt outperforms FWI and advanced DL methods in boundary delineation, structural preservation, noise robustness, and computational efficiency. These results highlight SC-UNeXt as a reliable tool for high-resolution seismic characterization of NGH reservoirs, thereby supporting sustainable exploration and risk assessment of marine hydrate resources.

©2026 China Geology Editorial Office.

1. Introduction

Natural gas hydrates (NGH) are crystalline compounds with a cage-like structure, consisting of natural gas and water molecule (Buffett BA, 2000; Boswell R and Collett TS, 2011). Given that these compounds remain stable exclusively under conditions of elevated pressure and low temperature (Gajanayake SM et al., 2023), they are predominantly found in deep-sea sediments or the permafrost regions of polar zones (Lu HL et al., 2007; Zhu YH et al., 2021). Furthermore, under standard temperature and pressure conditions, 1 m³ of hydrate can release approximately 164 m³ of natural gas (Ning FL et al., 2012; Gajanayake SM et al., 2023; Rao SH et al., 2025),

which makes NGH a promising candidate for the future utilization of energy resources. However, methane, the dominant component in most NGHs (Hester KC and Brewer PG, 2009; Yin ZY and Linga P, 2019; Wang P et al., 2024b), is a potent greenhouse gas, with a global warming potential substantially higher than that of carbon dioxide (Weissert H, 2000; Ye JL et al., 2018). Therefore, from both the perspectives of energy resource exploration and environmental sustainability, the precise identification and comprehensive characterization of NGH reservoirs are crucial (Li YL et al., 2023; Wei N et al., 2024).

Seismic inversion facilitates the development of high-resolution velocity models for NGH reservoirs (Wang JL et al., 2018; Kamal M et al., 2023; Yu H et al., 2024; Zhan LS et al., 2025), with advanced algorithms like Full Waveform inversion (FWI) providing a physics-based method for subsurface velocity reconstruction (Singh SC and Minshull TA, 1994; Satyavani N and Sain K, 2015; Djefall A et al., 2021). FWI exploits the sensitivity of pre-stack seismic data

First author: E-mail address: guozhenwei@csu.edu.cn (Zhen-wei Guo).

* Corresponding author: E-mail address: bochenwang29@csu.edu.cn (Bo-chen Wang).

Literary editor: Xi-jie Chen
 doi:10.31035/cg2025273

2096-5192/© 2026 China Geology Editorial Office.

to velocity variations and iteratively minimizes the misfit between synthetic and observed wavefields, thereby improving the identification of hydrate-bearing formations (Korenaga J et al., 1997; Wang et al., 2018). Despite its efficacy, FWI continues to encounter several significant limitations (Wang BF et al., 2022), including substantial computational requirements, non-uniqueness of solutions, and high sensitivity to the initial velocity model.

Consequently, to precisely delineate the boundaries of natural gas hydrate reservoirs, it is crucial to develop velocity inversion algorithms that ensure both high computational efficiency and reliable numerical stability. Deep learning (DL), regarded as a well-established approach in pattern recognition, has been extensively applied in computer vision (Simonyan K and Zisserman A, 2015; Ronneberger O et al., 2015; Liu Z et al., 2022) and geophysics (Li YS et al., 2022; Li CF et al., 2024; Ma K et al., 2025; Guo ZW et al., 2026). In seismic exploration, this technique has demonstrated its effectiveness in areas such as data denoising (Yang LQ et al., 2023; Chen JX et al., 2024; Gao ZQ et al., 2025), fault detection (Bönke W et al., 2024; Dou YM et al., 2024), and seismic facies classification (Noh K et al., 2023; Yang NX et al., 2024; Li M et al., 2025). Due to its efficiency, robustness, and high predictive accuracy, DL is progressively supplanting traditional seismic inversion methods. Nevertheless, the performance of DL-based inversion continues to be influenced by several key factors, including data quality (Yang YX et al., 2022; Li L et al., 2024; Wang P et al., 2024a), network architecture (Li FD et al., 2022; Wang HZ et al., 2023; Bustamante J et al., 2024), and the design of loss functions (Liu B et al., 2023; Zhang J et al., 2023; Zhang Y et al., 2023) in practical applications. Common enhancement strategies, such as architectural adjustments and loss-function optimization, improve model accuracy and efficiency, but often fail to capture boundary-sensitive features and deeper semantic information. This issue is particularly critical in NGH reservoirs (Liu Z et al., 2026; Wang W et al., 2026), where hydrate-bearing layers commonly overlie free-gas zones, forming high-low velocity thin-layer structures. Such configurations complicate boundary delineation and hinder reliable recovery of subtle stratigraphic features in complex seafloor settings.

To overcome the challenge of accurately inverting marine NGH reservoirs, this study proposes an efficient DL-based inversion network (SC-UNeXt). The main contributions of this work can be summarized as follows.

(i) A novel deep inversion architecture is proposed, which integrates a U-Net backbone with ConvNeXt residual blocks (CRB) to enhance feature extraction capability. Meanwhile, spatial-channel squeeze-and-excitation (scSE) attention mechanisms are embedded into the skip connections to improve the fusion of deep semantic features and shallow spatial details, thereby enhancing boundary delineation performance.

(ii) A hybrid loss function is designed by combining mean squared error (MSE) and multi-scale structural similarity

(MSSIM), enabling the network to jointly optimize pixel-wise accuracy and local structural consistency.

(iii) A multi-level perceptual difference loss is further introduced, which allows the network to capture high-level semantic information across multiple scales and better preserve global structural characteristics of NGH reservoirs.

To evaluate the proposed approach, this paper conducts extensive experiments on both NGH synthetic and the marine version of the 3D SEG/EAGE overthrust model with NGH (3DMO-NGH) datasets. The results confirm its superior inversion performance and ability to accurately delineate structural boundaries in NGH reservoirs, thereby supporting effective reservoir characterization.

2. Inversion framework with hybrid loss

To achieve high-fidelity seismic velocity inversion for natural gas hydrate (NGH) reservoirs, the proposed SC-UNeXt incorporates tailored network and hybrid loss functions into a unified, end-to-end inversion pipeline. The overall workflow, depicted in Fig. 1, delineates the sequential process from input data pre-processing to final inversion results, providing the structural basis for the detailed architectural and methodological descriptions that follow.

2.1. Network architecture

In this study, SC-UNeXt (Fig. 2a) adopts an encoder-decoder architecture based on U-Net and integrates three key components: CRB (Fig. 2b), pixel shuffle up-sampling (PSUS) module (Fig. 2c), and the scSE attention mechanism (Fig. 3). To refine information flow between encoder and decoder, scSE modules are embedded in the skip connections, enabling adaptive recalibration of features across both spatial and channel dimensions. The decoder reconstructs high-resolution outputs via PSUS, which effectively increases spatial resolution while reducing checkerboard artifacts. Together, these components improve the network's ability to extract multi-scale features, focus on salient regions, and produce high-fidelity inversion results with low computational overhead.

2.1.1. CRB-based downsampling module

Inspired by the architecture of vision transformers, the CRB (Fig. 2b) utilizes large 7×7 convolutional kernels to significantly expand the network's receptive field (Liu Z et al., 2022). At the same time, depth-wise separable convolutions are adopted to effectively reduce the number of parameters and computational cost. For normalization, the CRB replaces batch normalization (BN) commonly used in conventional convolutional neural networks with layer normalization (LN), which better mitigates gradient vanishing and exploding issues in deep architectures. Regarding activation functions, the CRB employs the gaussian error linear unit (GELU) in place of the rectified linear unit (ReLU). Unlike ReLU, which imposes a hard threshold, GELU provides a smoother and more continuous nonlinear

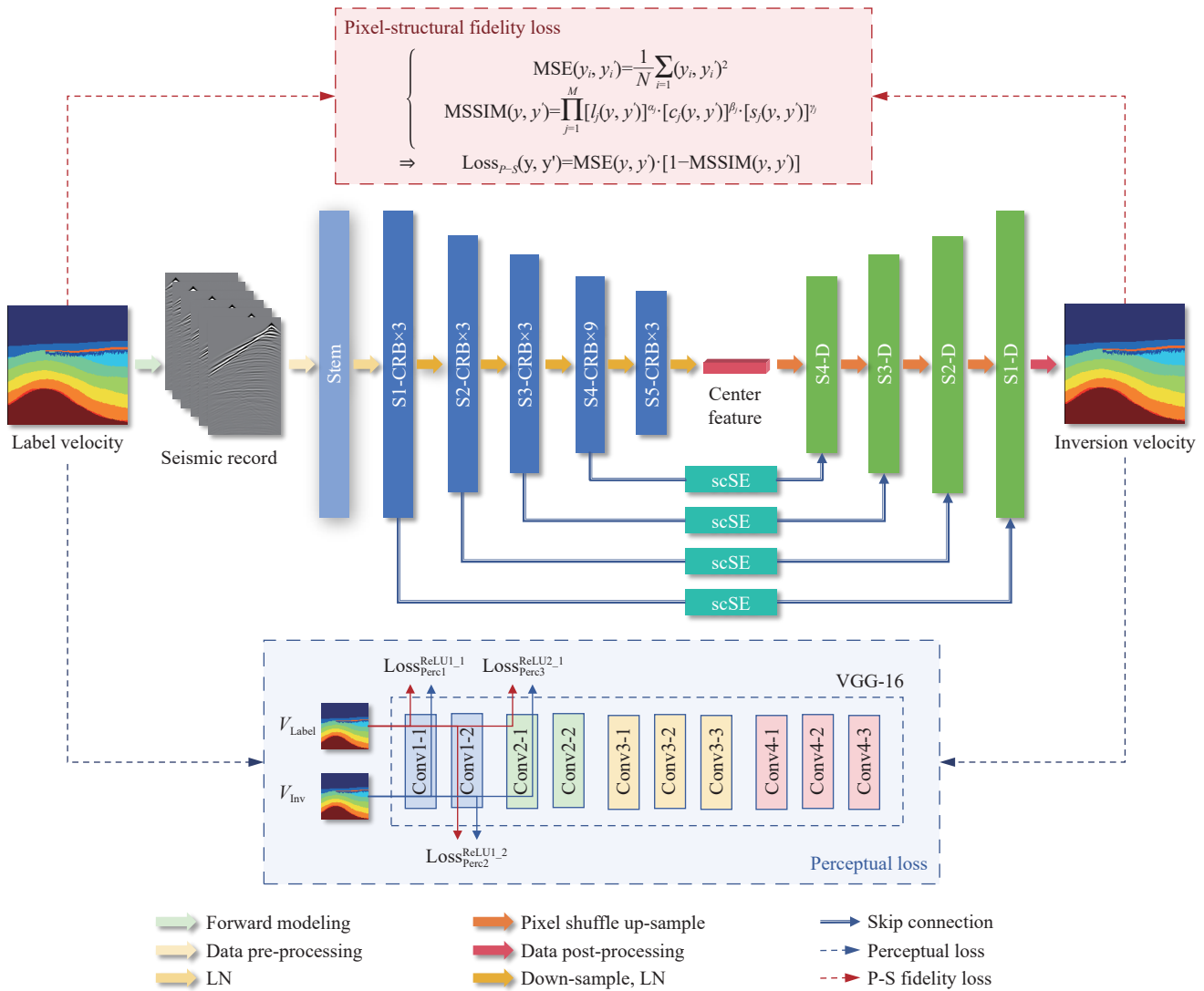


Fig. 1. Overall workflow of the SC-UNeXt with hybrid loss functions in an end-to-end pipeline. The hybrid loss consists of two components: The pixel-structural fidelity loss combining MSE and MSSIM (in red dashed box), and the perceptual loss (in blue dashed box).

response, thereby enhancing the model's fitting capacity and ability to model complex features.

2.1.2. PSUS module

To mitigate the checkerboard artifacts commonly associated with transposed convolution in the decoder, the PSUS module (Shi WZ et al., 2016) is adopted in this study network. This operation consists of three key steps: First, a conventional convolution operation expands the input to the channel dimension of feature maps, enriching feature representations for subsequent reorganization. Second, the PSUS operation efficiently redistributes high-frequency information from the channel dimension into the spatial domain, enabling structured up-sampling. At last, BN and ReLU activation are applied to standardize feature distributions and introduce non-linearity, thereby enhancing the network's ability to capture complex patterns (Fig. 2c).

2.1.3. scSE attention mechanism

In SC-UNeXt above, feature refinement within the skip connections is achieved by embedding the scSE attention

mechanism (Roy AG et al., 2018). The scSE attention module introduces two complementary attention branches (Fig. 3). The channel Squeeze-and-Excitation (cSE) branch applies global average pooling followed by a gating operation to assign channel-wise importance weights, thereby selectively amplifying informative channels while attenuating less relevant ones. In parallel, the spatial Squeeze-and-Excitation (sSE) branch employs pointwise (1×1) convolutions to generate a spatial attention map, highlighting salient regions within each feature map. The outputs of the two branches are fused to jointly recalibrate the feature responses along both spatial and channel dimensions. This dual-attention strategy enables SC-UNeXt to enhance discriminative seismic features, suppress background artifacts, and achieve more reliable velocity inversion in complex geological environments.

2.2. Loss function

2.2.1. Pixel-structure fidelity loss

As a typical regression task, inversion quantifies the

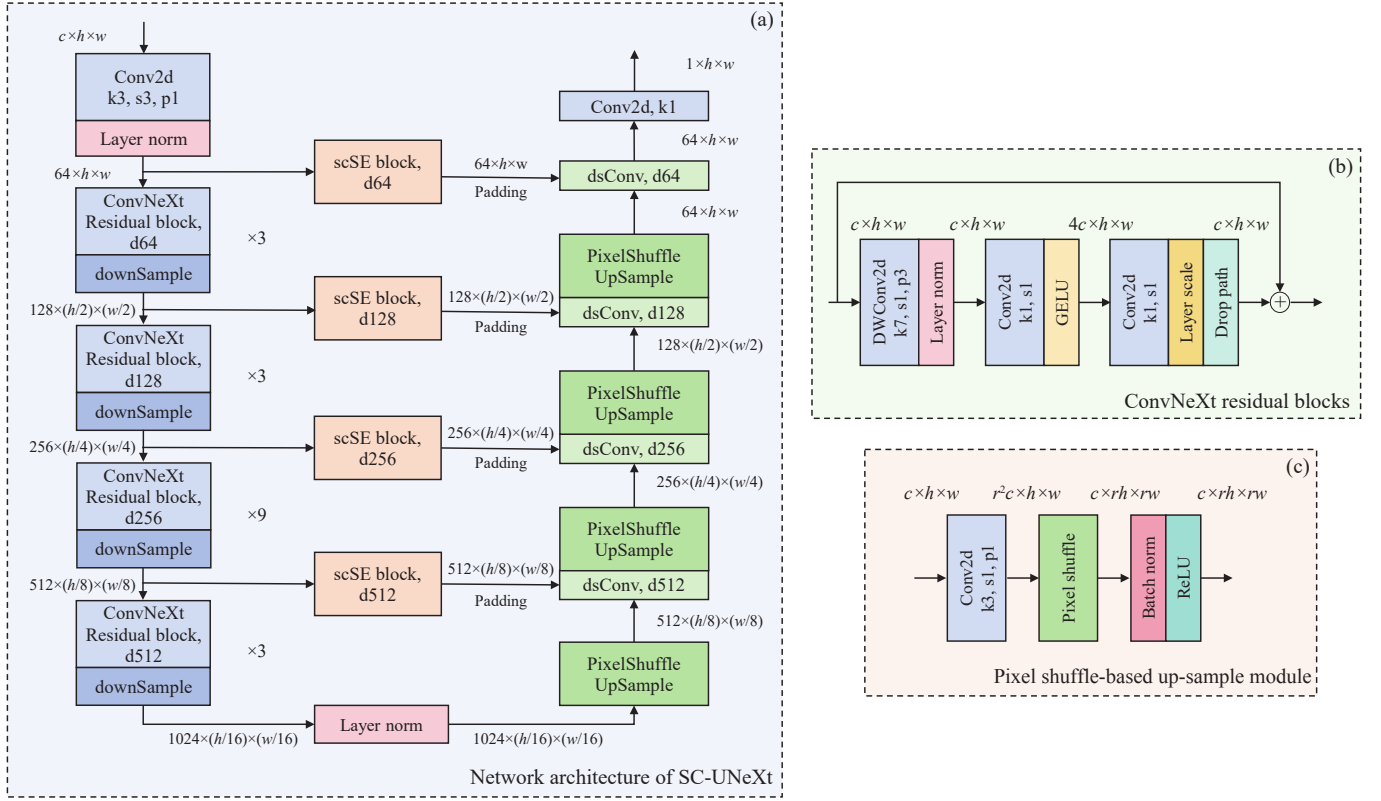


Fig. 2. Network architecture of SC-UNeXt with the detailed structure of its key components. a–Main structure of SC-UNeXt and its input/output dimension; b–structure of the CRB-based down-sampling module; c–structure of PSUS module. c , h , w represents the channel dimension, height, and width of the feature maps, respectively.

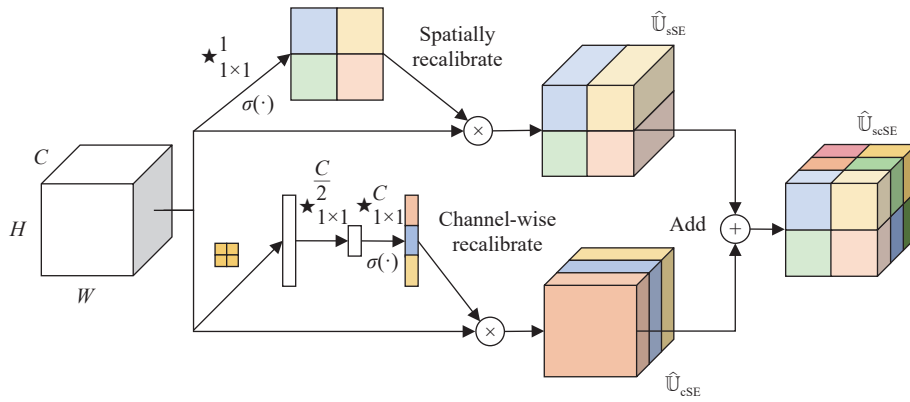


Fig. 3. Structure of scSE attention mechanism module.

discrepancy between inversion and true values by using the MSE loss function. The MSE measures the average squared difference between the label velocity and the inversion velocity defined as (Eq. 1):

$$\text{MSE}(y_i, y'_i) = \frac{1}{n} \sum_{i=1}^n (y'_i - y_i)^2 \quad (1)$$

where n denotes the total number of samples, y_i and y'_i represent the label and inversion velocity, respectively. However, the MSE primarily captures pixel-wise differences and is overly sensitive to outliers, thereby failing to account for local structural information such as textures and edges. To

address this limitation, the structural similarity index (SSIM) was introduced, which evaluates image quality by comparing luminance, contrast, and structural components, offering improved alignment with human visual perception. The SSIM between y_i and y'_i is defined as (Eq. 2):

$$\text{SSIM}(y, y') = \frac{(2\mu_y \mu_{y'} + C_1)(2\sigma_{yy'} + C_2)}{(\mu_y^2 + \mu_{y'}^2 + C_1)(\sigma_y^2 + \sigma_{y'}^2 + C_2)} \quad (2)$$

where μ_y and $\mu_{y'}$ are the local means, σ_y^2 and $\sigma_{y'}^2$ are the variances. $\sigma_{yy'}$ is the covariance between y and y' . C_1 and C_2 are constants to avoid instability when denominators are small. Building upon this, the multi-scale version computes SSIM across multiple resolutions to better capture perceptual

similarity at different spatial frequencies. The MSSIM is defined as (Eq. 3):

$$\text{MSSIM}(y, y') = \prod_{j=1}^M [l_j(y, y')]^{\alpha_j} \cdot [c_j(y, y')]^{\beta_j} \cdot [s_j(y, y')]^{\gamma_j} \quad (3)$$

In this equation, $l_j(y, y')$, $c_j(y, y')$ and $s_j(y, y')$ denote the luminance comparison, contrast comparison and structure comparison at scale j , respectively. α_j , β_j , γ_j are the weights assigned to the luminance, contrast, and structure components at scale j . M is the total number of scales.

To enable the network to focus on structural information of velocity during inversion and to mitigate the influence of extreme values on global optimization, this paper combines MAE and MSSIM to propose a pixel-structure (P-S) fidelity loss. The formulation of the loss is defined as (Eq. 4):

$$L_{\text{P-S}}(y, y') = \text{MSE}(y, y') \cdot (1 - \text{MSSIM}(y, y')) \quad (4)$$

The P-S fidelity loss balances numerical accuracy and structural fidelity by jointly minimizing prediction errors and preserving key spatial patterns in the data, enabling the model to produce outputs that are both quantitatively accurate and structurally representative of the input data.

2.2.2. Perceptual difference-based hybrid loss

To enhance the structural fidelity of seismic velocity inversion for NGH reservoirs, the authors incorporate a perceptual loss that evaluates similarity between the predicted and reference velocity models at the feature level. Unlike pixel-wise criteria, the perceptual loss exploits intermediate feature activations extracted from a pre-trained convolutional network (VGG-16), thereby embedding higher-level structural information into the optimization.

As illustrated in Fig. 4, the first column presents two velocity models inputted into the pre-trained VGG-16. The second, third and fourth column show feature maps extracted

from the Conv1_1, Conv1_2, and Conv2_1 layer of VGG-16, respectively. The feature maps from the shallow layers primarily capture low-level information, such as edges and shapes, which are essential for identifying basic structural features. In contrast, the deeper layers focus on higher-level abstractions, capturing more complex semantic information that corresponds to geological patterns and structures. This hierarchical feature extraction enables the perceptual loss to guide the inversion by preserving both low-level structural features and high-level geological semantics. Formally, let y and y' denote the reference and inversion velocity models, and let $\phi_s(\cdot)$ represent the feature map from the s -th layer of the pre-trained network ϕ . The perceptual loss is defined as (Eq. 5):

$$L_{\text{Perc}}(y, y') = \sum_{s \in S} \lambda_s \frac{1}{C_s H_s W_s} \|\phi_s(y') - \phi_s(y)\|_2^2 \quad (5)$$

where S is the set of all such selected layers, C_s , H_s and W_s correspond to the number of channels, height, and width of the feature maps at layer s , respectively. The weighting factor λ_s controls the relative contribution of each layer to the total perceptual loss. To ensure numerical accuracy, structural fidelity, and multi-level semantic coherence, the authors define a hybrid loss function that combines the P-S fidelity loss and the perceptual loss. The former enforces pixel-level and structural similarity, while the latter promotes feature-level alignment, enhancing the representation of geologically meaningful patterns. The hybrid loss is defined as (Eq. 6):

$$L_{\text{Hybrid}}(y, y') = L_{\text{P-S}}(y, y') + \lambda_{\text{Perc}} \cdot L_{\text{Perc}}(y, y') \quad (6)$$

where λ_{Perc} is the weighting factor that control the relative contributions of the perceptual loss. Here, λ_{Perc} is the weighting factor that is empirically determined based on the magnitude of different loss terms to balance their contributions during optimization.

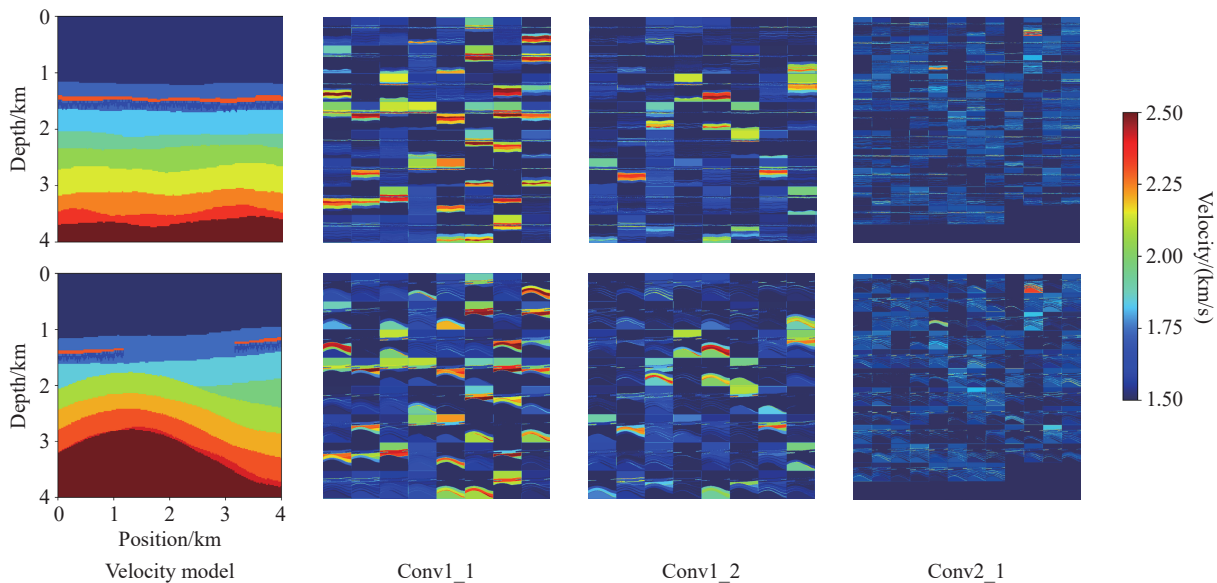


Fig. 4. Velocity models and its feature map examples extracted from VGG-16.

3. Data preparation and experimental setting

3.1. Data preparation

Two kinds of datasets are used in this study to evaluate the feasibility and effectiveness of the proposed method. The first is the synthetic dataset of marine NGHs and associated free gas, constructed based on the geological characteristics of a specific offshore area. The second dataset is derived from 3DMO-NGH, in which hydrate-bearing and free-gas layers were embedded through manual annotation within the gas hydrate stability zone (GHSZ), yielding a public dataset for testing the generalization capability of the proposed approach in layer inversion under realistic geological settings.

3.1.1. Synthetic dataset of marine NGH

To characterize the vertical P-wave velocity structure of NGH and free gas reservoirs, the authors employ a stochastic geological modeling approach (Wu XM et al., 2020; Ren YX et al., 2021). Reservoir modeling is performed using the parameters of gas hydrates and free gas specific to the study area, including velocity, thickness, and depth ranges constrained by well logs and production data (Ye JL et al., 2020; Xu Z et al., 2022). Based on downhole log data from the SHSC-4J1 and GMGS6-SH02 wells (Zhan LS et al., 2022), hydrate-bearing sediments occur within Unit I of the gas hydrate stability zone, located at 1495–1530 m in SHSC-4J1 and 1449–1495 m in GMGS6-SH02 (depths below sea level). Hydrate saturation was primarily determined from electrical resistivity logs using Archie's equation and further constrained through rock-physics modeling of sonic velocity data. Hydrate-bearing sediments with moderate to high hydrate saturation (generally exceeding 40% of pore space) may locally reach P-wave velocities approaching 2300 m/s in high-saturation intervals. This velocity increase is primarily attributed to the stiffening effect of hydrate acting as a load-bearing or cementing phase within the sediment framework. Therefore, a representative velocity of 2300 m/s is adopted in this study to characterize moderately to highly hydrate-saturated sediments in the synthetic models, representing an upper-endmember value for simplified parameterization rather than a strict average threshold. In contrast, free-gas-bearing sediments are characterized by significantly reduced P-wave velocities, because even a small amount of free gas can strongly decrease the bulk modulus of the sediment. Free-gas-bearing layers generally exhibit P-wave velocities in the range of 1500–1800 m/s. These intervals correspond to Unit III at 1545–1572 m in SHSC-4J1 and 1520–1539 m in GMGS6-SH02 (depths below sea level), equivalent to approximately 280–350 m beneath the seafloor after conversion using the local water depth. Accordingly, a velocity of 1600 m/s is assigned to the free-gas layer in the synthetic models.

To balance geological realism and computational efficiency, the synthetic velocity models are constructed using simplified layered structures that capture the essential velocity contrasts associated with seawater, background sediments, hydrate-bearing units, and underlying free-gas layers. The

model consists of 8–10 horizontal layers with smoothly varying interfaces generated using Gaussian perturbations. The upper seawater layer has a velocity randomly selected between 1500–1530 m/s and a thickness of 1000–1300 m, consistent with measurements from the target offshore region. Beneath this, sedimentary layers exhibit progressively increasing velocities, with random increments of 100–120 m/s and a cumulative thickness of 400–500 m. A hydrate-bearing unit, characterized by a representative velocity of 2300 m/s corresponding to moderately to highly hydrate-saturated sediments, is embedded approximately 200–500 m below the seafloor. When converted from depths reported below sea level in Zhan LS et al. (2022) using the local water depth (1226–1242 m), the hydrate-bearing intervals correspond to approximately 200–300 m beneath the seafloor; the extended range up to 500 m is introduced to explore broader geological variability within stochastic realizations, and the unit thickness varies between 60–200 m. A thin free-gas layer, assigned a representative lower velocity of 1600 m/s to account for the strong velocity-reducing effect of free gas, is inserted 20–60 m beneath the hydrate reservoir, consistent with the relative separation between the base of Unit I and the top of Unit III observed in the cited wells after converting depths from below sea level to below seafloor, with a thickness randomly assigned in the range of 40–100 m (Wang XJ et al., 2011). The velocity models (Fig. 5) are discretized on a two-dimensional grid of 200×200 points with a uniform spatial sampling interval of 20 m.

3.1.2. 3D SEG/EAGE marine overthrust model with NGH

The authors utilize the 3DMO-NGH to evaluate the generalization capability of this study method (Aminzadeh F, 1996). This model incorporates a 200 m-thick water layer and features velocity scaling to approximate sedimentary wave speeds typical of the North Sea environment. On this basis, the authors artificially embed a hydrate-bearing layer with a velocity of 2300 m/s and a thickness of 100–250 m within the theoretical gas hydrate stability zone (GHSZ), accompanied by an underlying free gas zone with a velocity of 1600 m/s and a thickness of 50–100 m.

Since the original dimensions of the 3DMO are 185×801×801 with a sampling interval of 25 m. To obtain a finer resolution, the authors interpolate along the x and z axes, resulting in a model of 200×800×801 with a 20 m sampling interval. The 3D data are subsequently sliced along the y -axis to generate 801 two-dimensional profiles. Based on the vertical extent of the z -axis, the authors further extract velocity models of size 200 × 200 by slicing along the x -axis with a step of 50. The overall workflow for dataset preparation is illustrated in Fig. 6. The velocity models (Fig. 7) are discretized on a two-dimensional grid of 200×200 points as same as the synthetic dataset. This adjustment ensures consistency of training parameters between this model and the synthetic dataset, thereby facilitating stable and reliable neural network training.

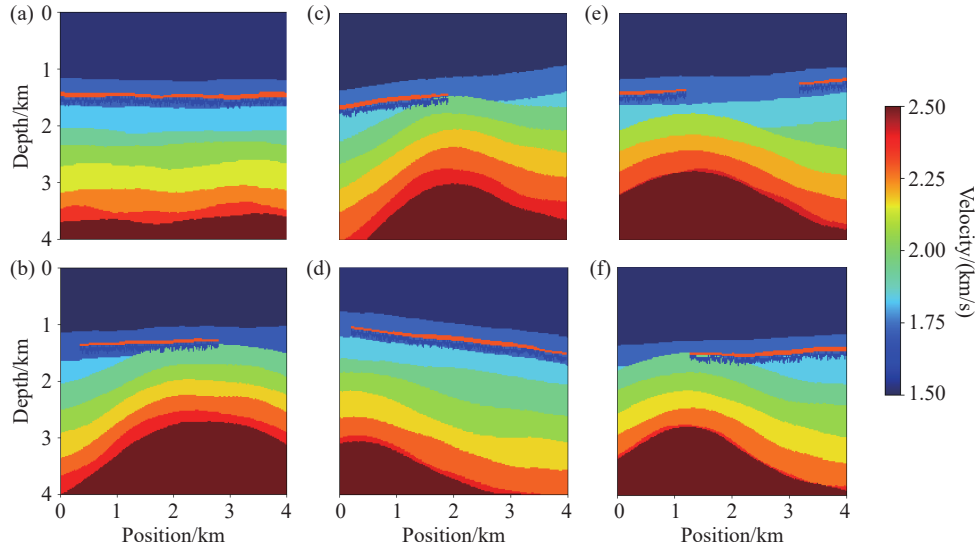


Fig. 5. Six representation examples of velocity models from synthetic dataset. The red geological target observed below the seafloor is interpreted as a gas hydrate-bearing layer, underlain by a blue geological target beneath that is indicative of a free gas zone. The surrounding features are characterized by laterally continuous, uniform stratified sedimentary structures.

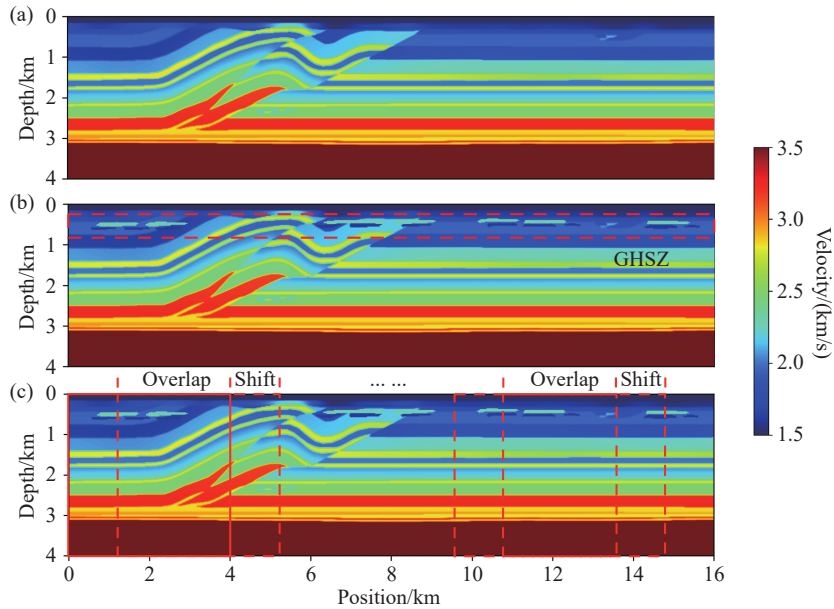


Fig. 6. Workflow for 3DMO-NGH dataset preparation. a–Slicing the interpolated 3D source model along the y -axis to obtain 801 two-dimensional velocity sections; b–incorporating a hydrate-bearing layer and its associated free gas zone within the theoretical GHSZ, aligned parallel to the seafloor; c–slicing along the x -axis according to the z -axis extent to generate velocity models for constructing the dataset used in SC-UNeXt training.

3.2. Forward modeling

To generate the training input, the authors simulate acoustic wave propagation using the second-order acoustic wave equation solved by the finite-difference time-domain method. The governing equation (Eq. 7) and source wavelet (Eq. 8) are given as:

$$\frac{1}{v^2(x, z)} \cdot \frac{\partial^2 u(x, z, t)}{\partial t^2} = \left(\frac{\partial^2 u(x, z, t)}{\partial x^2} + \frac{\partial^2 u(x, z, t)}{\partial z^2} \right) + s(t) \quad (7)$$

$$s(t) = [1 - 2(\pi f_m t)^2] e^{-(\pi f_m t)^2} \quad (8)$$

Here, (x, z) denote spatial coordinates and t represents propagation time. The function $v(x, z)$ refers to the P-wave velocity model, while $u(x, z, t)$ denotes the acoustic wavefield. The source term $s(t)$ is a ricker wavelet, and f_m is its dominant frequency. The wave equation is discretized using second-order central differences in both time and space. For the two datasets described in the previous section, the authors adopt consistent source and propagation parameters (Table 1). Under this configuration, synthetic seismic records are generated and subsequently employed as input to this study inversion network. Based on the above procedure, 3300 pairs of synthetic seismic records are constructed for both synthetic and 3DMO-NGH dataset.

3.3. Implementation details

Forward modeling, as well as the training, testing, and inversion of the neural network, are all implemented using PyTorch (Paszke A et al., 2019). The velocity models and their corresponding seismic records are paired and input into the network for training. Both datasets consist of 3300 pairs of data, randomly divided into two groups: The training set and the test set, containing 3000 and 300 pairs, respectively, without applying any geological or spatial partitioning rules. The hyper-parameters for training are provided in Table 2. To ensure a fair comparison and avoid potential bias arising from differences in training configurations, all DL models evaluated in this study were trained using the same optimizer, learning rate, learning rate scheduler, batch size, weight decay, and number of epochs. During training, the ReduceLROnPlateau scheduler was applied uniformly across all models to improve convergence. Specifically, this scheduler reduces the learning rate by a factor of 0.5 when the validation loss stops decreasing for 3 consecutive epochs. Due to the large size of the input data and hardware memory constraints, the batch size is empirically set to 5 to guarantee stable training.

3.4. Quantitative evaluation

In addition to directly comparing loss values, the authors

employ three quantitative metrics to evaluate the quality of the inversion velocity models: SSIM, peak signal-to-noise ratio (PSNR) and relative velocity error (RVE). The accuracy of the reconstructed velocity model is quantified by PSNR, which evaluates the similarity to the ground truth on a point-wise basis using a logarithmic scale. The equation is defined as follows (Eq. 9):

$$\text{PSNR}(y_i, y'_i) = 10 \cdot \log_{10} \left(\frac{\max(y_i)^2}{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2} \right) \quad (9)$$

where y_i and y'_i represent the label and predicted velocity, respectively, $\max(y_i)$ denotes the maximum possible velocity value in the label model, and N is the total number of grid points. The relative discrepancy between the label and predicted velocity values can be captured using the Relative Velocity Error (RVE), and is defined as (Eq. 10):

$$\text{RVE}(y_i, y'_i) = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - y'_i|}{y_i} \quad (10)$$

where y_i and y'_i represent the ground truth and predicted velocities at the i -th grid point, respectively. Collectively,

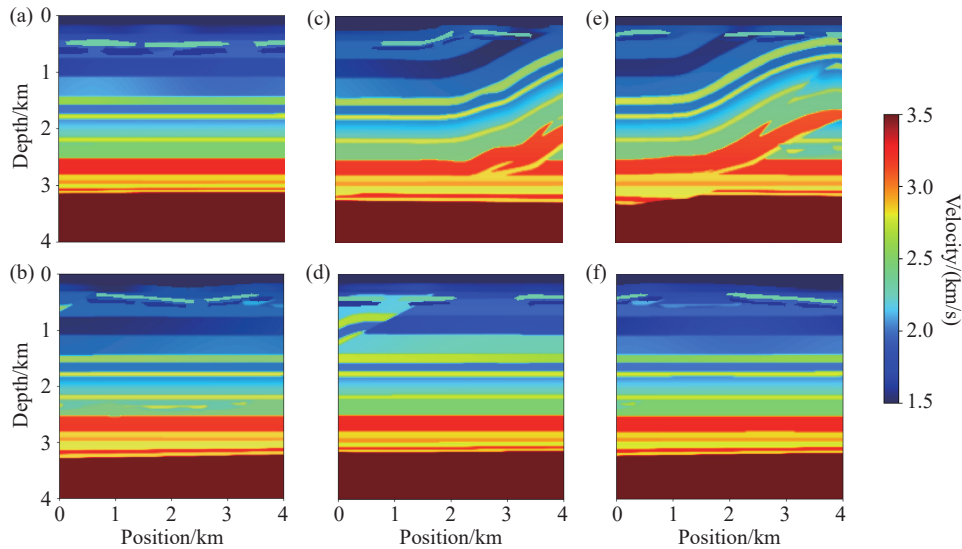


Fig. 7. Representative velocity sub-models extracted from the 3DMO-NGH. The light blue geological target observed below the seafloor is interpreted as a thin hydrate bearing layer parallel to the seafloor, underlain by a dark blue target indicative of a free gas zone parallel to the hydrate layer. The background comprises layered, overthrust, and fault structures inherent to the 3DMO model.

Table 1. Parameters for forward modeling using the finite-difference method (FDM).

Algorithm	Receiver interval	Time interval	Maximum travel time	Dominant frequency	Source points	Source interval	Boundary condition
FDM	20 m	0.001 s	3 s	15 Hz	6	720 m	PML

Notes: For each shot, all surface points are treated as receivers to achieve full-coverage seismic acquisition, resulting in a total of 200 receivers and a maximum offset of 4 km. To suppress boundary reflections, perfectly matched layers (PML) are applied as absorbing boundaries.

Table 2. Parameters for the training process of neutral networks, as determined through repeated experiments and empirical observations.

Optimizer	GPU	Learning rate	Weight decay	Batch size	Epoch
ADAM	RTX 4070	1×10^{-3}	1×10^{-4}	5	50

these metrics offer a robust assessment of inversion quality: SSIM measures the structural similarity between the predicted and ground truth models, PSNR reflects global fidelity, and RVE quantifies point-wise relative errors across the velocity field.

4. Performance evaluation and analysis

4.1. Assessment of inversion methods

4.1.1. Comparison with traditional baseline methods

To evaluate the effectiveness of the inversion method proposed in this study, the authors benchmark three representative methods: The conventional time-domain FWI, the widely used U-Net model in geophysical applications, and the SC-UNeXt method introduced in this work. The forward modeling settings for FWI adhere to the parameters listed in Table 1, with optimization performed using the L-BFGS method. Their respective training and inversion times are summarized in Table 3. Note that N/A indicates that time-domain FWI does not involve a training phase. To balance the number of training parameters between the two networks and to ensure the effectiveness of U-Net training, the epoch was set to 100, which is twice that of SC-UNeXt.

While time-domain FWI required approximately 240 minutes to complete the inversion for a single model, the DL-based approaches above required 144–162 minutes for the one-time training stage, after which each individual inversion during the testing stage was completed in about 20 seconds,

demonstrating markedly superior computational efficiency. It should be noted that time-domain FWI must be executed independently for each velocity model, whereas the DL-based methods benefit from pre-trained networks that can be stored and directly applied during the inference stage. In addition, FWI does not involve trainable parameters. The computational complexity of U-Net and SC-UNeXt mainly depends on their network architectures. SC-UNeXt introduces a more complex structure, resulting in slightly increased computational time compared to U-Net.

Furthermore, beyond computational efficiency, the data-driven approaches also demonstrate clear advantages over the physics-driven FWI in terms of inversion accuracy. As evidenced by the inversion results and residual maps (Figs. 8, 9), even when a relatively favorable initial model is provided, time-domain FWI exhibits noticeable discrepancies between the inverted and true velocities within the hydrate bearing and free gas layers, particularly near layer boundaries where the recovery is less satisfactory. These mismatches are further reflected by larger residuals concentrated around the interfaces between background sedimentary and hydrate layer. In contrast, the data-driven methods are capable of recovering the velocity distributions of hydrate and free gas bearing regions more accurately without relying on any initial velocity model. A closer inspection of the residual maps in the third and fourth columns of Fig. 9 reveals that, compared with SC-UNeXt, U-Net still suffers from localized velocity mismatches within the hydrate layer and the stratified background. Overall, when compared with both physics-

Table 3. Comparison of computational time consumption for time-domain FWI, U-Net and SC-UNeXt in traing and inversion processes on synthetic and 3DMO-NGH datasets.

Dataset	Process	FWI	U-Net	SC-UNeXt
Synthetic	Training	N/A	144 min 4 s	162 min 29 s
	Inversion	238 min 36 s	20 s	23 s
3DMO-NGH	Training	N/A	144 min 30 s	162 min 15 s
	Inversion	241 min 49 s	21 s	24 s

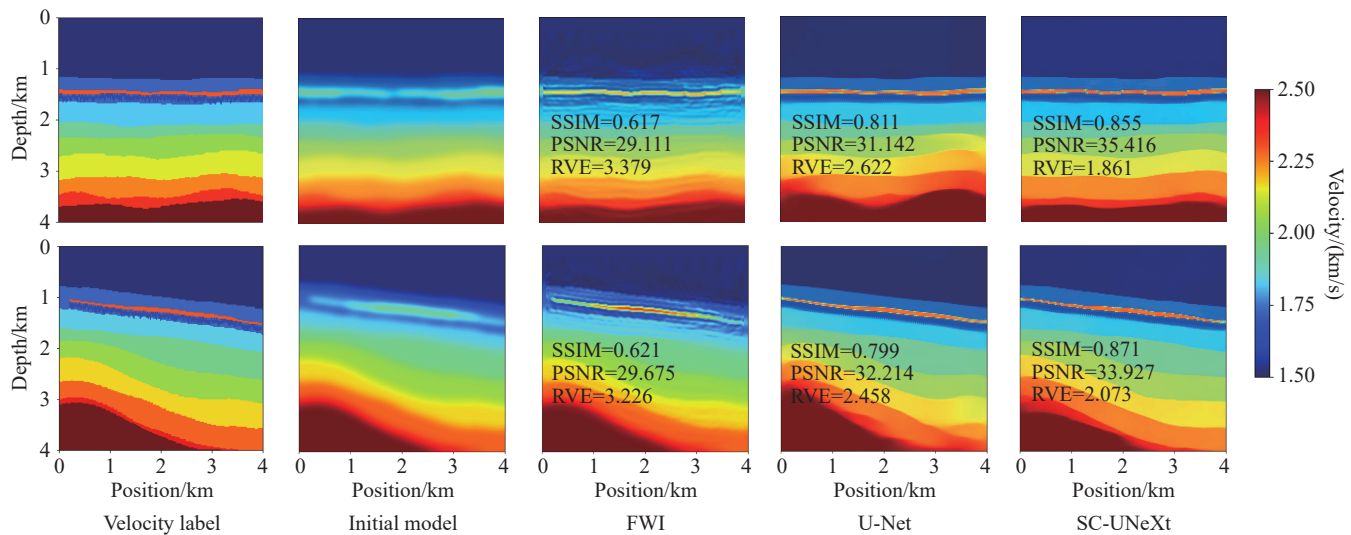


Fig. 8. Velocity models, initial models and inversion results of time-domain FWI, U-Net, and SC-UNeXt on the synthetic dataset.

driven and data-driven baseline methods, the proposed SC-UNeXt network achieves a favorable balance between computational efficiency and inversion accuracy. In order to comprehensively compare the metric trends of U-Net and SC-UNeXt during the training process, the authors present the metric evolution of both networks on test set of the synthetic dataset (Fig. 10). It can be observed that, under saturated training conditions, although U-Net exhibits a slightly faster convergence rate, SC-UNeXt consistently achieves superior and more stable performance across all four evaluation

metrics, demonstrating the robustness and overall effectiveness of the proposed framework.

4.1.2. Comparison with other networks

To further demonstrate the superiority of SC-UNeXt in hydrate velocity inversion, a comparison was made between the proposed network architecture and four widely used deep learning networks: DeepLabv3+, SwinUnet, TransUnet and ResUnet, representing convolutional and transformer-based networks in semantic segmentation. These networks were

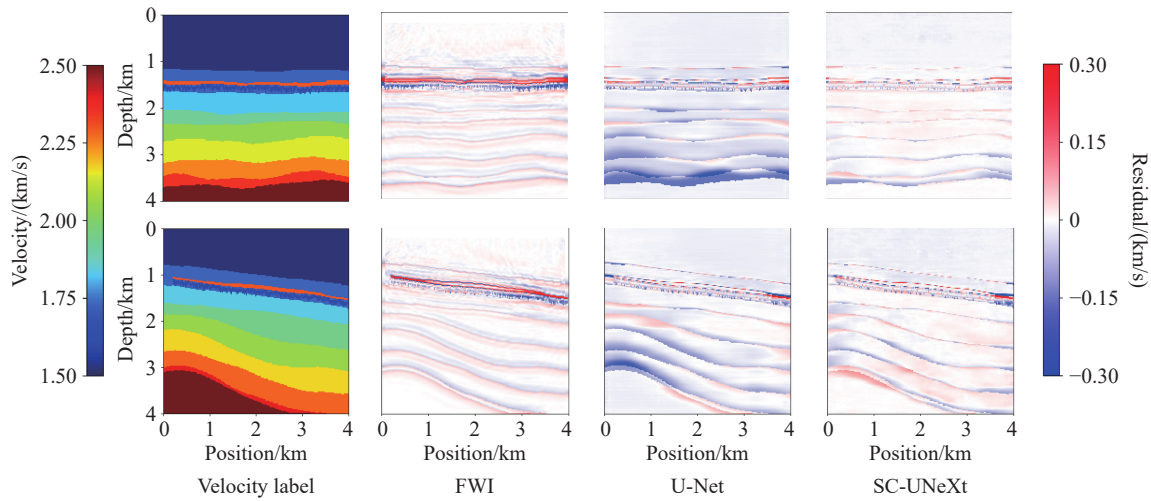


Fig. 9. Residuals of inversion results for time-domain FWI, U-Net, and SC-UNeXt on the synthetic dataset.

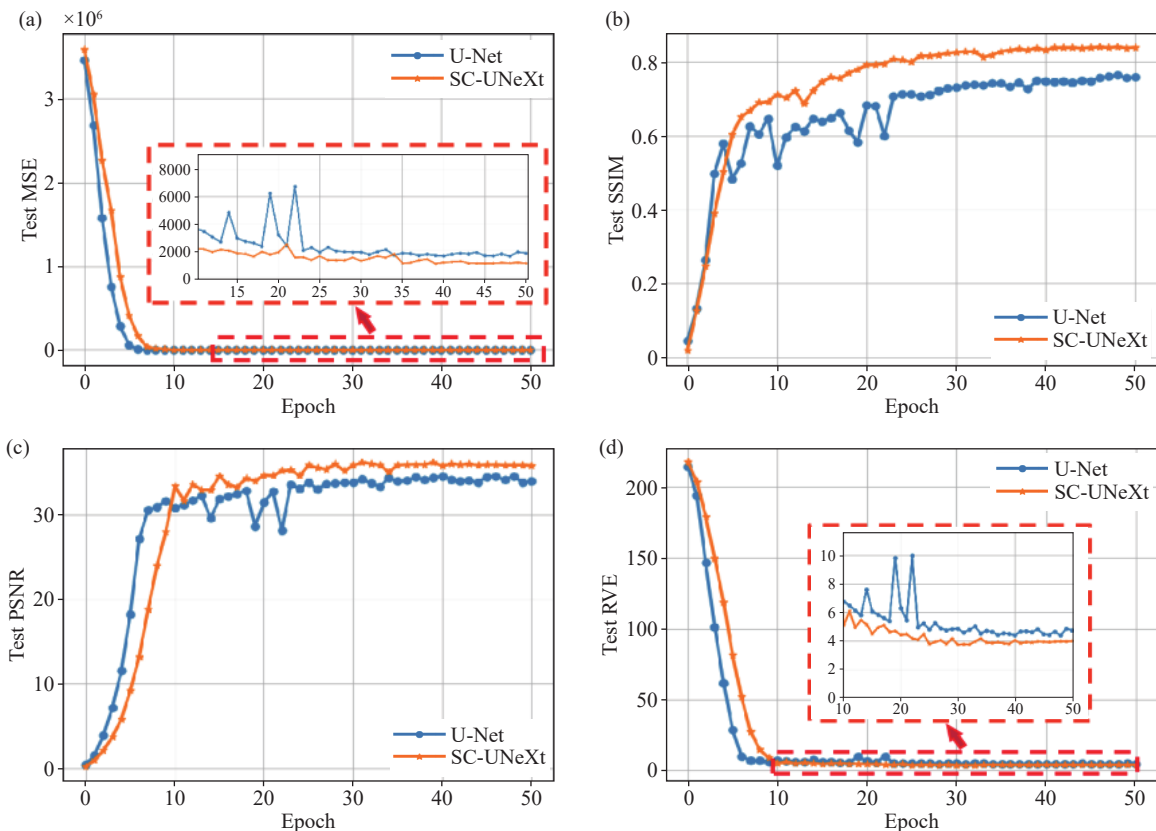


Fig. 10. Quantitative metrics history of both networks trained in synthetic test dataset. a–Test MSE; b–test SSIM; c–test PSNR; d–test RVE.

evaluated across two datasets, with the loss function set to MSE. The inversion performance of the networks is presented in the Fig. 11 and Fig. 12. From the inversion results, each network is generally able to predict the positions of the hydrate and free gas layers with relative accuracy. However, based on residual maps and evaluation metrics, the SC-UNeXt network not only captures more precise details of the hydrate layer but also provides a more accurate representation of the complex geological background. In both detailed structures and overall performance, the proposed network outperforms the others. The mean values of various evaluation metrics obtained by different networks on the synthetic datasets during the test period are summarized in Table 4.

As shown in Table 4, among the evaluated networks, SC-UNeXt demonstrates the most favorable trade-off between accuracy and efficiency, consistently outperforming the others in terms of structural fidelity and inversion quality. ResUnet also achieves competitive performance; however, SwinUnet

and TransUnet exhibit inferior results across most evaluation metrics. Notably, although SwinUnet and TransUnet require only about half the training time of the other networks, their transformer-based architectures focus on global contextual representation, resulting in lower effectiveness than convolutional-based networks in delineating geological boundaries and preserving fine-scale details. In particular, this shortcoming is most evident in the SSIM metric, where the discrepancy between the two categories of networks becomes especially pronounced in recovering subsurface structural features. Although the number of trainable parameters reflects model size, the training time is mainly determined by the network architecture and computational complexity rather than the parameter count alone.

4.2. Loss function optimization

To evaluate the effectiveness of the proposed loss function

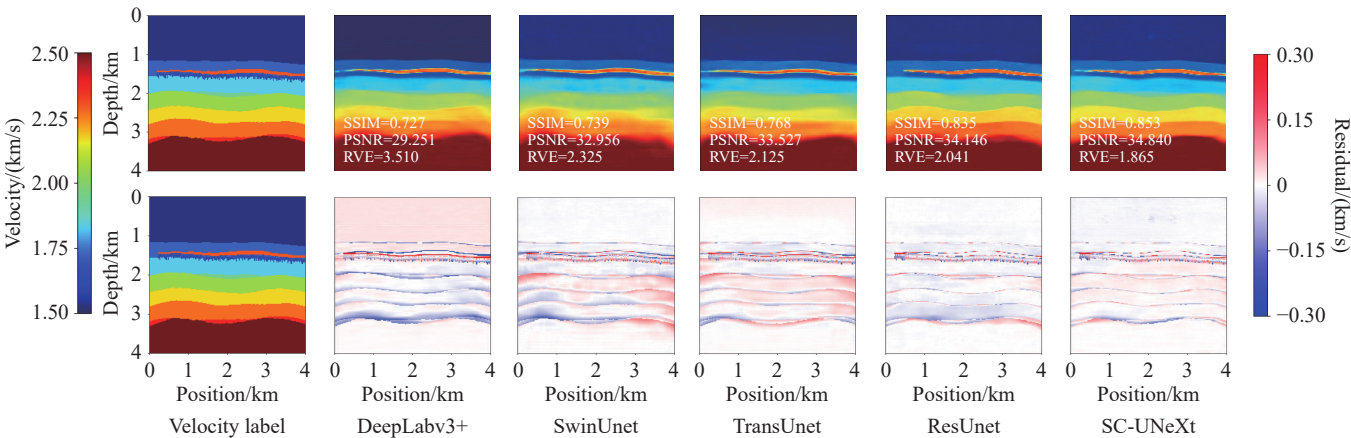


Fig. 11. Inversion results and residuals of DeepLabv3+, SwinUnet, TransUnet, ResUnet and SC-UNeXt on the synthetic dataset.

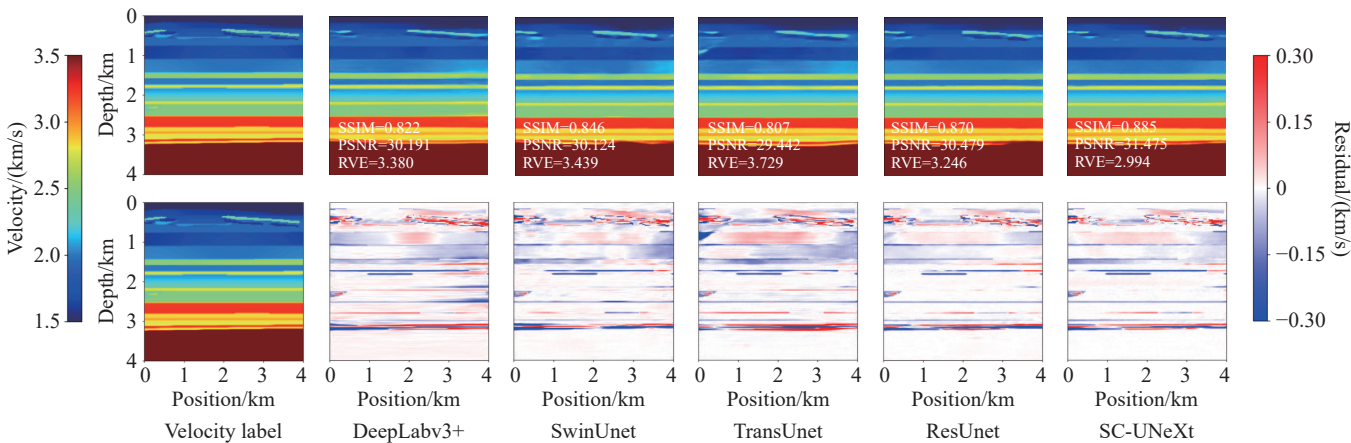


Fig. 12. Inversion results and residuals of DeepLabv3+, SwinUnet, TransUnet, ResUnet and SC-UNeXt on the 3DMO-NGH dataset.

Table 4. Performance, training time and model size comparison of five networks on the 3DMO-NGH dataset.

Dataset	Network	MSE	SSIM	PSNR	RVE	Time	Trainable parameters
Synthetic	DeepLabv3+	1463	0.767±0.064	34.895±4.197	4.299±1.606	85 min 59 s	67.04M
	SwinUnet	1633	0.717±0.086	34.397±4.310	4.469±1.823	75 min 52 s	68.29M
	TransUnet	1340	0.744±0.089	35.267±4.176	4.101±1.594	60 min 37 s	62.80M
	ResUnet	1255	0.832±0.090	35.573±3.549	3.987±1.411	162 min 36 s	77.23M
	SC-UNeXt	1091	0.841±0.085	36.182±3.281	3.758±1.425	153 min 25 s	75.26M

and to examine the role of its individual components in the optimization process, ablation experiments were conducted using four different losses on SC-UNet: MSE, a combination of MSE and MSSIM, a combination of MSE and perceptual loss, and the proposed hybrid loss function. The specific compositions of these loss functions are summarized in Table 5. Inversion experiments were performed on two independent datasets, and the reconstructed models as well as their residual distributions are illustrated in Figs. 13, 14.

The results demonstrate that while MSE alone achieves good numerical approximation, it has certain limitations in preserving global structures, leading to deviations in large-scale distributions. With the addition of MSSIM, the structural consistency of the inverted models improves significantly, although there is still room for improvement in fine-detail recovery, especially at the positions marked by the red arrows in the figures. Incorporating perceptual loss further

enhances the constraints in high-level feature space, enabling more accurate reconstruction of hydrate-bearing sediments and associated free-gas layers while improving semantic consistency across multiple scales. Ultimately, the proposed hybrid loss function achieves the optimal balance among numerical accuracy, structural fidelity, and semantic preservation, effectively reducing local residuals and substantially improving the overall inversion quality. From the corresponding regions in the first-row inversion results of Fig. 14, it can be observed that the hybrid loss function improves the delineation of the reservoir boundaries. The red boxes in the residual maps indicate the local differences associated with these regions.

4.3. Robustness and stability evaluation

To evaluate the stability and robustness of the proposed method, additional noise robustness experiments were

Table 5. Four different loss functions used in the ablation study.

Loss_1	Loss_2	Loss_3	Loss_4
MSE	MSE + (1-MSSIM)	MSE + $\lambda_{Perc} \cdot L_{Perc}$	MSE (1-MSSIM) + $\lambda_{Perc} \cdot L_{Perc}$

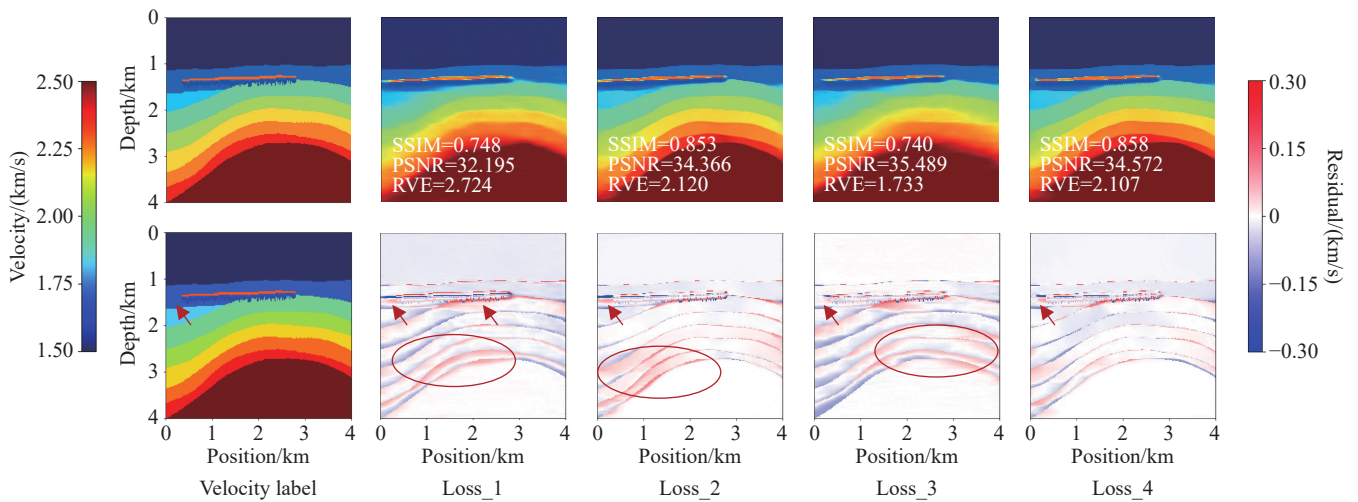


Fig. 13. Inversion results and residuals of different losses on the synthetic dataset.

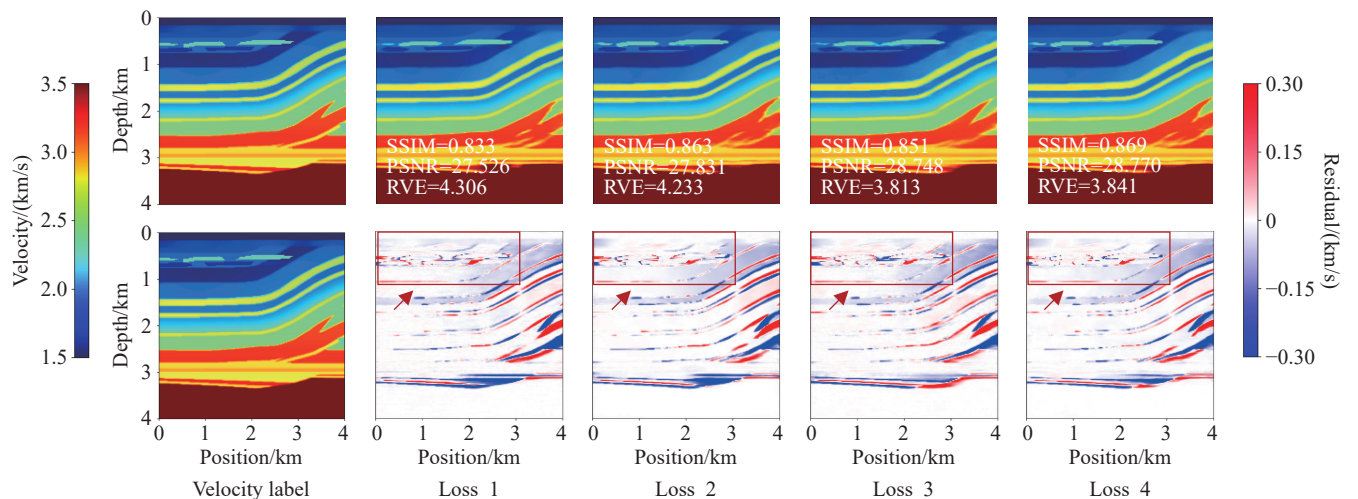


Fig. 14. Inversion results and residuals of different losses on the 3DMO-NGH dataset.

conducted on both the synthetic and 3DMO-NGH datasets. When seismic recordings were contaminated by varying levels of noise, the inversion results from SC-UNeXt largely maintained a high degree of consistency with the original velocity model (Fig. 15). Although the performance metrics for the noisy inversions were slightly lower than those obtained under noise-free conditions, the overall inversion quality remained stable. At a signal-to-noise ratio (SNR) of 15 decibel (dB) in the third column, the inversion successfully reconstructed the seabed stratigraphic background and accurately delineated the positions and structures of the hydrate and free-gas layers. As the noise level increased (SNR ≤ 10 dB and SNR ≤ 5 dB), the inversion results showed varying degrees of deviation compared to the low-noise or noise-free scenarios, but the hydrate and free-gas layers were still effectively represented in the velocity model. To further assess the robustness of the method, the authors also performed inversions on the noise-contaminated 3DMO-NGH dataset. Fig. 16 demonstrates similar patterns to those

observed with the synthetic dataset, indicating that SC-UNeXt retains a high level of inversion performance even under noisy conditions. The experimental results show that the proposed method can generate stable and reliable velocity models from noise-contaminated data, demonstrating strong noise resilience and robustness.

5. Discussion

Building upon the results presented in the previous chapter, it is evident that the proposed method (Fig. 2) provides stable and accurate solutions for velocity inversion of marine NGH and their associated free-gas thin layers. The superior performance of SC-UNeXt can be attributed to three key innovations. First, the incorporation of CRB modules in the encoder substantially enhances the extraction of deep seismic features. Second, embedding the scSE attention mechanism within skip connections enables more effective fusion of spatial and channel-wise information. Third, replacing conventional linear interpolation or transposed

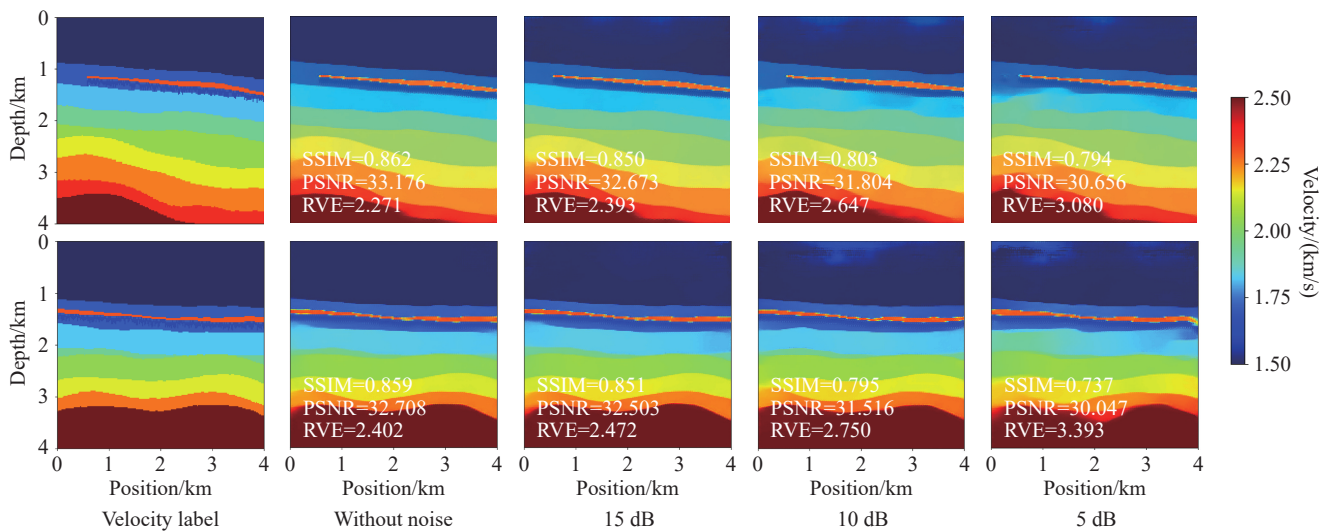


Fig. 15. Noise sensitivity analysis of the proposed method in synthetic datasets.

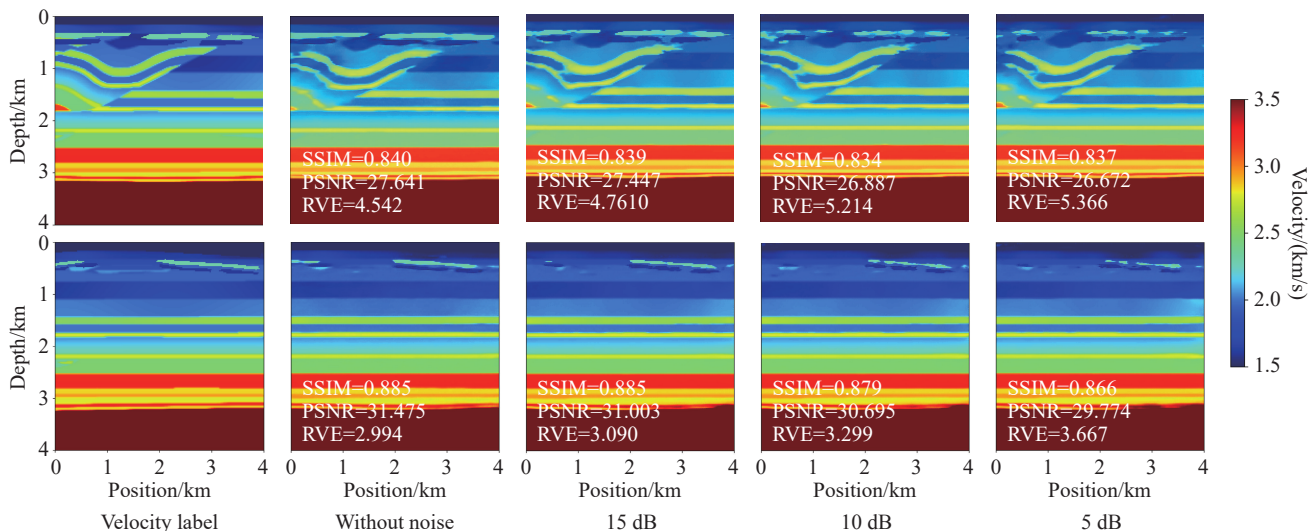


Fig. 16. Noise sensitivity analysis of the proposed method in 3DMO-NGH datasets.

convolution with PSUS in the decoder improves up-sampling quality and mitigates artifacts during feature map reconstruction.

Compared with other U-Net-based architectures, the CRB in SC-UNeXt not only enlarge the receptive field through large convolutional kernels but also maintain training efficiency owing to their lightweight design. At the same time, the scSE module, consisting of complementary sSE and cSE branches, effectively emphasizes critical regions such as hydrate layers and free-gas boundaries. The scSE mechanism further redistributes high-frequency details from the channel dimension to the spatial domain, thereby achieving structured upsampling. Taken together, the quantitative assessments in Table 4 and the visual comparisons in Fig. 11 and Fig. 12 demonstrate that SC-UNeXt can accurately characterize anomalous velocity structures of hydrate and free-gas layers in complex marine settings, while maintaining computational efficiency.

Beyond architectural improvements, the design of the loss function plays a pivotal role in capturing thin-layer features. To this end, a hybrid loss function was introduced. Conventional error-sensitive losses (MSE) primarily capture pixel-level discrepancies but fail to account for spatial correlations among geological structures. To overcome this limitation, MSSIM was incorporated to capture multi-scale structural similarities, thereby strengthening spatial correlation modeling. In addition, perceptual loss was employed. Unlike direct pixel-wise or structural comparisons, perceptual loss leverages a pre-trained feature extractor to compare feature maps at multiple levels, highlighting the boundaries, distribution, and relative positioning of geological targets. This approach enables deep fusion of semantic information across scales and improves consistency in feature space. As illustrated in Fig. 13 and Fig. 14, the hybrid loss not only enhances the delineation accuracy of hydrate and free-gas boundaries, but also alleviates the non-uniqueness inherent in inversion problems.

Finally, the evaluation on both synthetic data and the 3DMO-NGH dataset underscores the strong generalization ability, robustness to noise, and adaptability of SC-UNeXt to complex real-world scenarios. Nevertheless, field data typically contain stronger noise and are limited in sample size, which poses greater challenges to DL-based inversion methods. In particular, (1) the geological settings encountered in practice are often more diverse than those represented in the training data, and variations in stratigraphic architecture, faulting, lithological heterogeneity, or acquisition conditions may lead to a noticeable domain shift, thereby degrading the model performance when transferred to unseen areas. (2) The current workflow still relies heavily on synthetic data for supervision, which may introduce a mismatch between synthetic and field data due to simplified physics, incomplete noise modeling, and mismatches in source–receiver geometry and bandwidth, limiting the robustness of the learned mapping. (3) Moreover, uncertainties in field labels and the non-uniqueness of inversion further complicate the

interpretation of predicted results.

To mitigate these limitations, several strategies can be considered, including incorporating more realistic forward modeling with broader geological priors, enriching training samples through noise-aware data augmentation, and employing transfer learning or domain adaptation to better align synthetic and field data distributions. In particular, physics-guided loss functions could explicitly incorporate forward modeling constraints (wave equation residuals or data-consistency terms) into the optimization objective, thereby reducing the discrepancy between synthetic and field data distributions and alleviating domain shift by enforcing physically consistent predictions. Such constraints may regularize the solution space and improve robustness when the geological settings deviate from those represented in the training data. In addition, semi-supervised or self-supervised learning frameworks that exploit unlabeled field data may alleviate label scarcity, while ensemble-based uncertainty quantification can provide confidence measures for inversion outputs. In light of recent advances, physics-guided data-driven approaches appear particularly promising (Lin YZ et al., 2023; Zhang J et al., 2024), as they demonstrate potential advantages under small-sample conditions and contribute to improved interpretability (Li P et al., 2024). Within the current SC-UNeXt framework, such physics-guided strategies could be implemented by augmenting the hybrid loss with additional physics-consistency terms, without altering the network architecture, making integration technically feasible. Consequently, future research should focus on integrating physical priors into the network optimization process and incorporating physics-constrained loss functions to further enhance training stability and the reliability of inversion results. Moreover, transfer learning is feasible within the present framework by initializing the network with weights pre-trained on large-scale synthetic datasets and subsequently fine-tuning selected layers using limited field data. Freezing low-level feature extractors while adapting higher-level layers could help preserve general seismic feature representations while accommodating region-specific geological variations, thereby improving adaptability to new survey areas.

6. Conclusions

This study aims to improve the accuracy and robustness of seismic velocity inversion for marine natural gas hydrate and free-gas systems, by developing a deep learning-based end-to-end inversion framework that directly maps seismic data to velocity models. The main contributions of this work can be summarized as follows.

(i) This study establishes two novel synthetic seismic velocity datasets grounded in realistic geological settings. The first dataset represents a layered subsurface model, while the second, the 3DMO-NGH dataset, captures complex marine geological structures with detailed distributions of natural gas hydrate and free-gas layers. These datasets provide a reliable benchmark for evaluating DL-based seismic velocity

inversion methods under both simple and highly heterogeneous subsurface conditions.

(ii) To accurately delineate the spatial distribution and boundaries of NGH and free-gas reservoirs, the authors propose SC-UNeXt, an efficient end-to-end framework for direct seismic velocity inversion. The framework integrates a tailored encoder–decoder architecture with CRB-based downsampling, scSE attention-enhanced skip connections, and PSUS-based upsampling. This design enables adaptive feature extraction across multiple scales and enhances sensitivity to velocity anomalies associated with thin hydrate and free-gas layers, leading to accurate and stable inversion results.

(iii) A hybrid loss function is further introduced to jointly enforce pixel-level accuracy, structural continuity, and feature-level semantic consistency. By combining P-S fidelity loss with perceptual loss, the proposed optimization strategy significantly improves the reconstruction of complex marine velocity fields and the resolution of thin geological layers. Extensive experiments demonstrate that SC-UNeXt achieves strong robustness to noise and favorable generalization performance while maintaining computational efficiency.

The proposed method is particularly suitable for marine NGH reservoirs with complex stratigraphic structures, provided that sufficient training data are available and the noise level is within the evaluated range ($\text{SNR} \geq 5$ dB), where stable inversion performance has been observed. Nevertheless, challenges remain in applying DL-based inversion to field data, particularly regarding noise contamination and limited labeled samples. Future work will focus on incorporating physics-guided priors and constraint-aware loss functions to enhance interpretability, training stability, and practical applicability in real marine environments.

CRedit authorship contribution statement

Zhen-wei Guo, Jun Lei and Bo-chen Wang conceived of the presented idea. Jun Lei carried out the experiment. All authors discussed the results and contributed to the final manuscript.

Declaration of competing interest

The authors declare no conflicts of interests.

Acknowledgment

This research was jointly supported by the Natural Science Foundation of China (42574195).

References

Aminzadeh F. 1996. 3-D salt and overthrust seismic models. In: Weimer P, Davis TL. Applications of 3-D seismic data to exploration and production. Tulsa, American Association of Petroleum Geologists, 376–391. doi: [10.1306/St42605C29](https://doi.org/10.1306/St42605C29).

Bönke W, Alaei B, Torabi A, Oikonomou D. 2024. Data augmentation

for 3D seismic fault interpretation using deep learning. Marine and Petroleum Geology, 162, 106706. doi: [10.1016/j.marpetgeo.2024.106706](https://doi.org/10.1016/j.marpetgeo.2024.106706).

Boswell R, Collett TS. 2011. Current perspectives on gas hydrate resources. Energy & Environmental Science, 4(4), 1206–1215. doi: [10.1039/C0EE00203H](https://doi.org/10.1039/C0EE00203H).

Buffett BA. 2000. Clathrate hydrates. Annual Review of Earth and Planetary Sciences, 28(1), 477–507. doi: [10.1146/annurev.earth.28.1.477](https://doi.org/10.1146/annurev.earth.28.1.477).

Bustamante J, Fabien-Ouellet G, Duchesne MJ, Ibrahim A. 2024. Deep-learning viscoelastic seismic inversion for mapping subsea permafrost. Geophysics, 89(4), 339–353. doi: [10.1190/geo2022-0759.1](https://doi.org/10.1190/geo2022-0759.1).

Chen JX, Chen GX, Li J, Du RS, Qi YL, Li CF, Wang NJ. 2024. Efficient seismic data denoising via deep learning with improved MCA-SCUNet. IEEE Transactions on Geoscience and Remote Sensing, 62, 1–14. doi: [10.1109/TGRS.2024.3355972](https://doi.org/10.1109/TGRS.2024.3355972).

Djeffal A, Pecher IA, Singh SC, Crutchley GJ, Kaipio J. 2021. Improved imaging of gas hydrate reservoirs and their plumbing system using 2D elastic full-waveform inversion. Interpretation, 9(3), 955–968. doi: [10.1190/INT-2020-0163.1](https://doi.org/10.1190/INT-2020-0163.1).

Dou YM, Li KW, Dong MH, Xiao Y. 2024. FaultSSL: Seismic fault detection via semisupervised learning. Geophysics, 89(3), 79–91. doi: [10.1190/geo2023-0550.1](https://doi.org/10.1190/geo2023-0550.1).

Gajanayake SM, Gamage RP, Li XS, Huppert H. 2023. Natural gas hydrates—Insights into a paradigm-shifting energy resource. Energy Reviews, 2(1), 100013. doi: [10.1016/j.enrev.2024.100089](https://doi.org/10.1016/j.enrev.2024.100089).

Gao ZQ, Yang LS, Yang Z, Gao JH. 2025. Physics-constrained deep learning for attenuation compensation of nonstationary seismic data. IEEE Transactions on Geoscience and Remote Sensing, 63. doi: [10.1109/TGRS.2025.3541495](https://doi.org/10.1109/TGRS.2025.3541495).

Guo ZW, Lu KX, Chen YJ. 2026. A semi-airborne transient electromagnetic data denoising method based on TOC-VMD and improved threshold. Journal of Geophysics and Engineering, 23(5), 1541–1555. doi: [10.1093/jge/gxag066](https://doi.org/10.1093/jge/gxag066).

Hester KC, Brewer PG. 2009. Clathrate hydrates in nature. Annual review of marine science, 1(1), 303–327. doi: [10.1146/annurev.marine.010908.163824](https://doi.org/10.1146/annurev.marine.010908.163824).

Kamal M, Ali A, Amin Y, Farid A, Sarosh B. 2023. Advanced data processing as a tool to enhance the vertical and horizontal resolution of bottom simulating reflector (BSR): Implications for GAS hydrate exploitation. Pure and Applied Geophysics, 180(6), 2369–2387. doi: [10.1007/s00024-023-03257-3](https://doi.org/10.1007/s00024-023-03257-3).

Korenaga J, Holbrook W, Singh S, Minshull T. 1997. Natural gas hydrates on the southeast US margin: Constraints from full waveform and travel time inversions of wide-angle seismic data. Journal of Geophysical Research: Solid Earth, 102(B7), 15345–15365. doi: [10.1029/97JB00725](https://doi.org/10.1029/97JB00725).

Li CD, Feng PF, Jiang XH, Zhang S, Meng J, Li BC. 2024. Extensive identification of landslide boundaries using remote sensing images and deep learning method. China Geology, 7(2), 277–290. doi: [10.31035/cg2023148](https://doi.org/10.31035/cg2023148).

Li FD, Guo ZW, Pan XP, Liu JX, Wang YY, Gao DW. 2022. Deep learning with adaptive attention for seismic velocity inversion. Remote Sensing, 14(15), 3810. doi: [10.3390/rs14153810](https://doi.org/10.3390/rs14153810).

Li L, Zeng XB, Pan XP, Peng L, Tan YY, Liu JX. 2024. Microseismic velocity inversion based on deep learning and data augmentation. Applied Sciences, 14(5), 2194. doi: [10.3390/app14052194](https://doi.org/10.3390/app14052194).

- Li M, Yan XS, Wu QH. 2025. A self-supervised deep learning framework for seismic facies segmentation. *Expert Systems with Applications*, 288. doi: [10.1016/j.eswa.2025.128290](https://doi.org/10.1016/j.eswa.2025.128290).
- Li P, Liu ML, Alfarrarj M, Tahmasebi P, Grana D. 2024. Probabilistic physics-informed neural network for seismic petrophysical inversion. *Geophysics*, 89(2), 17–32. doi: [10.1190/geo2023-0214.1](https://doi.org/10.1190/geo2023-0214.1).
- Li YL, Hu QB, Wu NY, Wang HB, Sun XF, Hu GW, Sun ZW, Jiang YJ. 2023. Acoustic characterization for creep behaviors of marine sandy hydrate-bearing sediment. *Scientific Reports*, 13(1), 22199. doi: [10.1038/s41598-023-49523-1](https://doi.org/10.1038/s41598-023-49523-1).
- Li YS, Peng C, Ran XJ, Xue LF, Chai SL. 2022. Soil geochemical prospecting prediction method based on deep convolutional neural networks-Taking Daqiao Gold Deposit in Gansu Province, China as an example. *China Geology*, 5(1), 71–83. doi: [10.31035/cg2021044](https://doi.org/10.31035/cg2021044).
- Lin YZ, Theiler J, Wohlberg B. 2023. Physics-guided data-driven seismic inversion: Recent progress and future opportunities in full-waveform inversion. *IEEE Signal Processing Magazine*, 40(1), 115–133. doi: [10.1109/MSP.2022.3217658](https://doi.org/10.1109/MSP.2022.3217658).
- Liu B, Jiang P, Wang QY, Ren YX, Yang SL, Cohn AG. 2023. Physics-driven self-supervised learning system for seismic velocity inversion. *Geophysics*, 88(2), 145–161. doi: [10.1190/geo2021-0302.1](https://doi.org/10.1190/geo2021-0302.1).
- Liu Z, Li YL, Ye RJ, Dong L, Qi MH, Wang W, Li ZK, Ning FL, Wu NY. 2026. A hybrid unsupervised-to-supervised machine learning framework for fracture segmentation in natural gas hydrate-bearing sediments. *Engineering Geology*, 364, 108624. doi: [10.1016/j.enggeo.2026.108624](https://doi.org/10.1016/j.enggeo.2026.108624).
- Liu Z, Mao HZ, Wu CY, Feichtenhofer C, Darrell T, Xie SN. 2022. A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11976–11986. doi: [10.1109/CVPR52688.2022.01167](https://doi.org/10.1109/CVPR52688.2022.01167).
- Lu HL, Seo YT, Lee JW, Moudrakovski I, Ripmeester JA, Chapman NR, Coffin RB, Gardner G, Pohlman J. 2007. Complex gas hydrate from the Cascadia margin. *Nature*, 445(7125), 303–306. doi: [10.1038/nature05463](https://doi.org/10.1038/nature05463).
- Ma K, Liu JJ, Lu SQ, Huang ZH, Tian M, Deng JY, Xie Z, Qiu QJ. 2025. Improved SE-UNet network-based semantic segmentation and implied geological significance from geological map. *China Geology*, 8, 1–19. doi: [10.31035/cg2023146](https://doi.org/10.31035/cg2023146).
- Ning FL, Yu YB, Kjelstrup S, Vlugt TJH, Glavatskiy K. 2012. Mechanical properties of clathrate hydrates: status and perspectives. *Energy & Environmental Science*, 5(5), 6779–6795. doi: [10.1039/C2EE03435B](https://doi.org/10.1039/C2EE03435B).
- Noh K, Kim D, Byun J. 2023. Explainable deep learning for supervised seismic facies classification using intrinsic method. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1–11. doi: [10.1109/TGRS.2023.3236500](https://doi.org/10.1109/TGRS.2023.3236500).
- Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin ZM, Gimelshein N, Antiga L, Desmaison A, Kopf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai JJ, Chintala S. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32. doi: [10.48550/arXiv.1912.01703](https://doi.org/10.48550/arXiv.1912.01703).
- Rao SH, Li ZC, Lu HL, Deng YJ. 2025. A review of gas hydrate formation characteristics at interfaces. *Fuel*, 392, 134863. doi: [10.1016/j.fuel.2025.134863](https://doi.org/10.1016/j.fuel.2025.134863).
- Ren YX, Nie LC, Yang SL, Jiang P, Chen YK. 2021. Building complex seismic velocity models for deep learning inversion. *IEEE Access*, 9, 63767–63778. doi: [10.1109/ACCESS.2021.3051159](https://doi.org/10.1109/ACCESS.2021.3051159).
- Ronneberger O, Fischer P, Brox T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-assisted Intervention*, 234–241. doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- Roy AG, Navab N, Wachinger C. 2018. Concurrent spatial and channel ‘squeeze & excitation’ in fully convolutional networks. In *International Conference on Medical Image Computing and Computer-assisted Intervention*, 421–429. doi: [10.1007/978-3-030-00928-1_48](https://doi.org/10.1007/978-3-030-00928-1_48).
- Satyavani N, Sain K. 2015. Seismic insights into bottom simulating reflection (BSR) in the Krishna-Godavari basin, Eastern Margin of India. *Marine Georesources and Geotechnology*, 33(3), 191–201. doi: [10.1080/1064119X.2013.797059](https://doi.org/10.1080/1064119X.2013.797059).
- Shi WZ, Caballero J, Huszar F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang ZH. 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1874–1883. doi: [10.1109/CVPR.2016.207](https://doi.org/10.1109/CVPR.2016.207).
- Simonyan K, Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations*. doi: [10.48550/arXiv.1409.1556](https://doi.org/10.48550/arXiv.1409.1556).
- Singh SC, Minshall, TA. 1994. Velocity structure of a gas hydrate reflector at Ocean Drilling Program site 889 from a global seismic waveform inversion. *Journal of Geophysical Research: Solid Earth*, 99(B12), 24221–24233. doi: [10.1029/94JB02168](https://doi.org/10.1029/94JB02168).
- Wang BF, Li JK, Wang P, Si WP, Wang ZK. 2022. Visco-acoustic full waveform inversion using decoupled fractional Laplacian constant-Q wave equation and optimal transport-based misfit function. *Exploration Geophysics*, 53(2), 140–150. doi: [10.1080/08123985.2021.1901573](https://doi.org/10.1080/08123985.2021.1901573).
- Wang HZ, Lin J, Dong XT, Lu SP, Li Y, Yang BJ. 2023. Seismic velocity inversion transformer. *Geophysics*, 88(4), R513–R533. doi: [10.1190/geo2022-0283.1](https://doi.org/10.1190/geo2022-0283.1).
- Wang JL, Jaiswal P, Haines SS, Hart PE, Wu SG. 2018. Gas hydrate quantification using full-waveform inversion of sparse ocean-bottom seismic data: A case study from Green Canyon Block 955, Gulf of Mexico. *Geophysics*, 83(4), B167–B181. doi: [10.1190/geo2017-0414.1](https://doi.org/10.1190/geo2017-0414.1).
- Wang P, Cui YA, Zhou L, Li JY, Pan XP, Sun Y, Liu JX. 2024a. Multi-task learning for seismic elastic parameter inversion with the lateral constraint of angle-gather difference. *Petroleum Science*, 21(6), 4001–4009. doi: [10.1016/j.petsci.2024.06.010](https://doi.org/10.1016/j.petsci.2024.06.010).
- Wang P, Xu CK, Yin ZY, Song SX, Xu CJ, Dai SD. 2024b. A DEM-based generic modeling framework for hydrate-bearing sediments. *Computers and Geotechnics*, 171, 106287. doi: [10.1016/j.compgeo.2024.106287](https://doi.org/10.1016/j.compgeo.2024.106287).
- Wang W, Wu NY, Wang, L., Wang LZ, Liu Z, Li YL. 2026. The interaction of gas hydrate with submarine landslides studied with cascading machine learning. *Landslides*, 23, 1187–1200. doi: [10.1007/s10346-025-02687-5](https://doi.org/10.1007/s10346-025-02687-5).
- Wang XJ, Wu SG, Lee M, Guo YQ, Yang SX, Liang JQ. 2011. Gas hydrate saturation from acoustic impedance and resistivity logs in the Shenhu area, South China Sea. *Marine and Petroleum Geology*, 28(9), 1625–1633. doi: [10.1016/j.marpetgeo.2011.07.002](https://doi.org/10.1016/j.marpetgeo.2011.07.002).
- Wei N, Pei J, Li HT, Zhou SW, Zhao JZ, Kvamme B, Coffin RB, Zhang LH, Zhang Y, Xue J. 2024. Classification of natural gas hydrate resources: Review, application and prospect. *Gas Science and Engineering*, 124, 205269. doi: [10.1016/j.jgsce.2024.205269](https://doi.org/10.1016/j.jgsce.2024.205269).

- Weissert H. 2000. Deciphering methane's fingerprint. *Nature*, 406(6794), 356–357. doi: [10.1038/35019230](https://doi.org/10.1038/35019230).
- Wu XM, Geng ZC, Shi YZ, Pham N, Fomel S, Caumon G. 2020. Building realistic structure models to train convolutional neural networks for seismic structural interpretation. *Geophysics*, 85(4), 1–48. doi: [10.1190/geo2019-0375.1](https://doi.org/10.1190/geo2019-0375.1).
- Xu Z, Hu T, Pang XQ, Wang EZ, Liu XH, Wu ZY, Chen D, Li CR, Zhang XW, Wang T. 2022. Research progress and challenges of natural gas hydrate resource evaluation in the South China Sea. *Petroleum Science*, 19(1), 13–25. doi: [10.1016/j.petsci.2021.12.007](https://doi.org/10.1016/j.petsci.2021.12.007).
- Yang LQ, Fomel S, Wang SD, Chen XH, Chen W, Saad OM, Chen YK. 2023. Denoising of distributed acoustic sensing data using supervised deep learning. *Geophysics*, 88(1), 91–104. doi: [10.1190/geo2022-0138.1](https://doi.org/10.1190/geo2022-0138.1).
- Yang NX, Li GF, Li TH, Zhao DF, Gu WW. 2024. An improved deep dilated convolutional neural network for seismic facies interpretation. *Petroleum Science*, 21(3), 1569–1583. doi: [10.1016/j.petsci.2023.11.027](https://doi.org/10.1016/j.petsci.2023.11.027).
- Yang YX, Zhang XT, Guan Q, Lin YZ. 2022. Making invisible visible: Data-driven seismic inversion with spatio-temporally constrained data augmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–16. doi: [10.1109/TGRS.2022.3144636](https://doi.org/10.1109/TGRS.2022.3144636).
- Ye JL, Qin XW, Qiu HJ, Liang QY, Dong YF, Wei JG, Lu HL, Lu JA, Shi YH, Zhong C, Xia Z. 2018. Preliminary results of environmental monitoring of the natural gas hydrate production test in the South China Sea. *China Geology*, 1(2), 202–209. doi: [10.31035/cg2018029](https://doi.org/10.31035/cg2018029).
- Ye JL, Qin XW, Xie WW, Lu HL, Ma BJ, Qiu HJ, Liang JQ, Lu JA, Kuang ZG, Lu C, Liang QY, Wei SP, Yu YJ, Liu CS, Li B, Shen KX, Shi HX, Lu QP, Li J, Kou BB, Song G, Li B, Zhang HE, Lu HF, Ma C, Dong YF, Bian H. 2020. The second natural gas hydrate production test in the South China Sea. *China Geology*, 3(2), 197–209. doi: [10.31035/cg2020043](https://doi.org/10.31035/cg2020043).
- Yin ZY, Linga P. 2019. Methane hydrates: A future clean energy resource. *Chinese Journal of Chemical Engineering*, 27(9), 2026–2036. doi: [10.1016/j.cjche.2019.01.005](https://doi.org/10.1016/j.cjche.2019.01.005).
- Yu H, Wang J, Lei ZS, Kuang ZG, Deng W, Lu L. 2024. Characterization of natural gas hydrate constrained by well and seismic data in Qiongdongnan Basin. *Energies*, 17(21), 5411. doi: [10.3390/en17215411](https://doi.org/10.3390/en17215411).
- Zhan LS, Kang DJ, Lu HL, Lu JA. 2022. Characterization of coexistence of gas hydrate and free gas using sonic logging data in the Shenhu Area, South China Sea. *Journal of Natural Gas Science and Engineering*, 101, 104540. doi: [10.1016/j.jngse.2022.104540](https://doi.org/10.1016/j.jngse.2022.104540).
- Zhan LS, Matsushima, J, Liu HT, Lu HL. 2025. Evaluation and modeling of velocity dispersion and frequency-dependent attenuation in gas hydrate-bearing sediments. *Marine and Petroleum Geology*, 171, 107204. doi: [10.1016/j.marpetgeo.2024.107204](https://doi.org/10.1016/j.marpetgeo.2024.107204).
- Zhang J, Sun H, Zhang G, Huang XG, Han L. 2024. Simultaneous physics and model-guided seismic inversion based on deep learning. *IEEE Transactions on Geoscience and Remote Sensing*, 99, 1–11. doi: [10.1109/TGRS.2024.3443970](https://doi.org/10.1109/TGRS.2024.3443970).
- Zhang J, Zhao XY, Chen YK, Sun H. 2023. Domain knowledge-guided data-driven prestack seismic inversion using deep learning. *Geophysics*, 88(2), 31–47. doi: [10.1190/geo2021-0560.1](https://doi.org/10.1190/geo2021-0560.1).
- Zhang Y, Meng DC, Zhou YF, Song LW, Dong HL. 2023. Seismic velocity inversion based on physically constrained neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–17. doi: [10.1109/TGRS.2023.3339783](https://doi.org/10.1109/TGRS.2023.3339783).
- Zhu YH, Pang SJ, Xiao R, Zhang S, Lu ZQ. 2021. Natural gas hydrates in the Qinghai-Tibet Plateau: Characteristics, formation, and evolution. *China Geology*, 4(1), 17–31. doi: [10.31035/cg2021025](https://doi.org/10.31035/cg2021025).