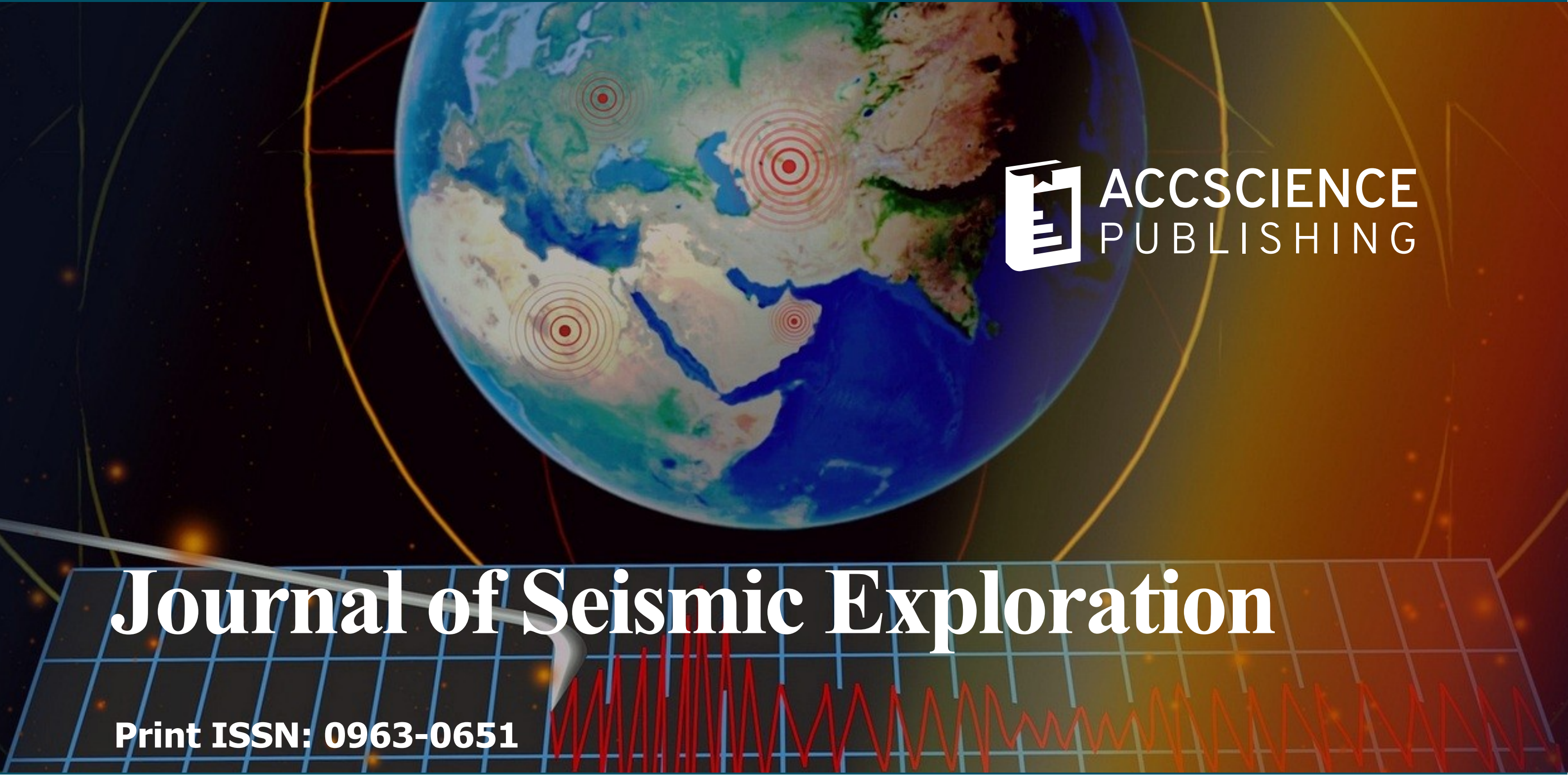
Journal of
Seismic Exploration

ISSN: 0963-0651

Volume 34 · Issue 5
2025

ACCSCIENCE
PUBLISHING

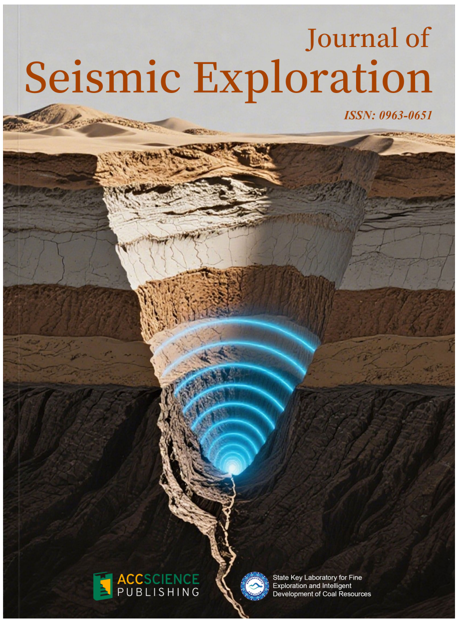
State Key Laboratory for Fine
Exploration and Intelligent
Development of Coal Resources



Journal of Seismic Exploration

Print ISSN: 0963-0651

The Journal of Seismic Exploration (JSE) is an international medium for the publication of research in seismic modeling, processing, inversion, interpretation, field techniques, borehole techniques, tomography, instrumentation and software. The aim and goal of JSE is to serve the interests and objectives of seismic petroleum exploration. All seismic processing methods have assumptions and prerequisites- and when they are satisfied methods are effective-and when they are violated methods become ineffective and can fail. The latter contributes to drilling dry hole exploration wells and suboptimal development wells.



Journal of
Seismic Exploration
ISSN: 0963-0651

ACCSCIENCE
PUBLISHING

State Key Laboratory for Fine
Exploration and Intelligent
Development of Coal Resources

About the Publisher

AccScience Publishing is a publishing company based in Singapore. We publish a range of high-quality, open-access, peer-reviewed journals and books from a broad spectrum of disciplines.

Contact Us

Managing Editor
jse.office@accscience.sg

AccScience Publishing
9 Raffles Place, Republic Plaza 1 #06-00 Singapore 048619.

Volume 34 • Issue 5 • 2025
ISSN 0963-0651 (print)

Journal of Seismic Exploration

ACCSCIENCE
PUBLISHING

Access Science Without Barriers

ACCSCIENCE
PUBLISHING

Full issue copyright © 2025 AccScience Publishing
All rights reserved. Without permission in writing from the publisher, this full issue publication in its entirety may not be reproduced or transmitted for commercial purposes in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system. Permissions may be sought from jse.office@accscience.sg

Article copyright © Respective Author(s)
See articles for copyright year. All articles in this full issue publication are open-access. There are no restrictions in the distribution and reproduction of individual articles, provided the original work is properly cited. However, permission to reuse copyrighted materials of an article for commercial purposes is applicable if the article is licensed under Creative Commons Attribution-NonCommercial License. Check the specific license before reusing.

JOURNAL OF SEISMIC EXPLORATION

ISSN: 0963-0651 (print)

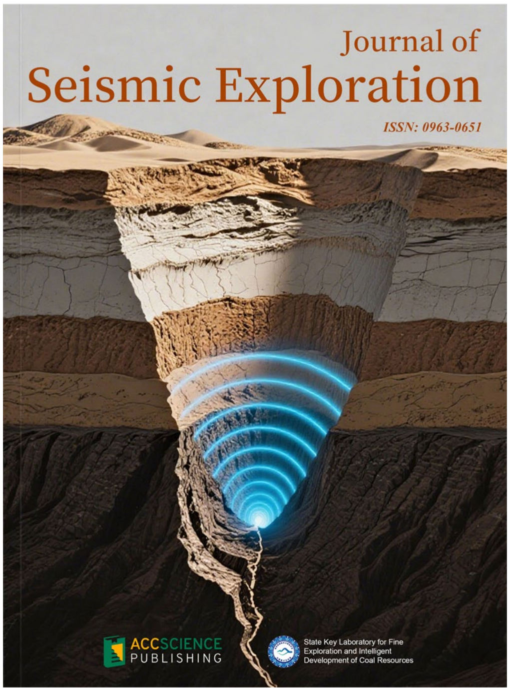
Editorial and Production Credits

Publisher: AccScience Publishing
Managing Editor: Gary Liu
Production Editor: Lim Mingyuan
Article Layout and Typeset: Sinjore Technologies (India)

For all advertising queries, contact jse.office@accscience.sg

Supplementary file

Supplementary files of articles can be obtained at https://accscience.com/journal/JSE/34/5.



Disclaimer

AccScience Publishing is not liable to the statements, perspectives, and opinions contained in the publications. The appearance of advertisements in the journal shall not be construed as a warranty, endorsement, or approval of the products or services advertised and/or the safety thereof. AccScience Publishing disclaims responsibility for any injury to persons or property resulting from any ideas or products referred to in the publications or advertisements. AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Seismic Exploration

Editorial Board

Editors-in-Chief

Executive Editor-in-Chief

*Editorial Board Members**

Early Career Editorial Board

ACCSCIENCE PUBLISHING

Volume 34 I Issue 5 I 2025 I Page 1 – 98 I

Journal of Seismic Exploration

https://accscience.com/journal/JSE

CONTENTS

ARTICLES

1 **A fluid factor inversion method using the frequency-domain two-step sub-band regularization**
Peng Zhang, Ying Xiao, Peng Xiao, Pang Chen, Wangyang Xu

18 **Application of the spatial autocorrelation method incorporating damping optimization for identifying goafs in cultural heritage sites**
Jiale Liu, Xiaodong Wang, Yanhai Liu, Ke Ren, Shangqing Zhang

36 **GMLAN: Grouped-residual and multi-scale large-kernel attention network for seismic image super-resolution**
Anxin Zhang, Zhenbo Guo, Shiqi Dong, Zhiqi Wei

53 **Dual-branch dense network for seismic background noise elimination**
Wei Wang, Haoliang Chen, Dekuan Chang, Xinyang Wang, Shujiang Wang, Dong Li

66 **Transposed arrangement strategy-based deep learning for seismic data crossline interpolation**
Jiyun Yu, Yonghwan Joo, Daeung Yoon

81 **Hybrid convolutional neural network–graph attention network–gradient boosting decision tree model for seismic impedance inversion prediction**
Tianwen Zhao, Guoqing Chen, Cong Pang, Palakorn Seenoi, Nipada Papukdee, Piyapatr Busababodhin, Yiru Du

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

A fluid factor inversion method using the frequency-domain two-step sub-band regularization

Peng Zhang[1], **Ying Xiao**[1], **Peng Xiao**[2,3]*, **Pang Chen**[1], **and Wangyang Xu**[3]

[1]Geophysical R&D Institute, Geophysical-COSL, China Oilfield Services Ltd., Tianjin, China

[2]Sichuan Energy Investment Group Co., Ltd., Chengdu, Sichuan, China

[3]Reservoir Geophysics Laboratory, School of Geoscience and Technology, Southwest Petroleum University, Chengdu, Sichuan, China

Abstract

In oil and gas seismic exploration, fluid factors are key parameters for identifying reservoir fluid properties and evaluating reservoir potential. Although regularization methods are commonly used to enhance inversion stability, traditional time–domain lateral constraint methods struggle to effectively address the issue of abrupt lateral stratigraphic variations. This paper aims to improve the prediction accuracy and stability of fluid factors in such scenarios. Based on the characteristic that frequency-domain seismic data exhibit stronger correlation during stratigraphic abrupt changes, this paper proposes a frequency-domain two-step sub-band regularization inversion method. First, a difference operator is introduced into the frequency domain to construct the objective function, and the Alternating Direction Method of Multipliers algorithm is adopted as the solution. Furthermore, a two-step sub-band regularization strategy is proposed: First, integrating low-frequency and high-frequency residual terms into the same objective function; then performing inversion in two stages. These stages first use low-frequency data inversion to obtain the general profile of the subsurface structure as an initial model, followed by using high-frequency data inversion based on this to obtain detailed information. Theoretical model tests have verified the superiority of this method under complex geological conditions. Field application in practical work areas shows that the frequency-domain two-step method significantly outperforms the traditional time–domain lateral constraint L1 regularization method in terms of vertical resolution and lateral continuity. This method provides a more accurate and stable solution for fluid identification under complex geological conditions.

Keywords: Frequency-domain regularization; Fluid factor; Lateral constraint; Alternating Direction Method of Multipliers; Pre-stack inversion

***Corresponding author:**
Peng Xiao
(15680819559@163.com)

Citation: Zhang P, Xiao Y, Xiao P, Chen P, Xu W. A fluid factor inversion method using the frequency-domain two-step sub-band regularization. *J Seismic Explor*. 2025;34(5):1-17. doi: 10.36922/JSE025310048

Received: July 28, 2025

1st revised: August 22, 2025

2nd revised: September 1, 2025

3rd revised: September 16, 2025

4th revised: September 23, 2025

Accepted: October 11, 2025

1. Introduction

The accurate identification of subsurface fluids is crucial for oil and gas exploration and development. It enables geologists and engineers to gain a deeper understanding of reservoir characteristics, predict the distribution of oil and gas reservoirs, and thereby improve the success rate of exploration and development efficiency. As a result, research on fluid factors has become a key and dynamic research direction. Initially, fluid factors were parameters constructed based on variations in P-wave and S-wave velocities, used to detect hydrocarbon anomalies in reservoirs. Over time, researchers have not only developed multi-parameter methods and innovated direct extraction techniques, but also clarified the petrophysical significance of fluid factors and applied them to the exploration of unconventional oil and gas resources. The concept of fluid factor was first proposed by Smith and Gidlow[1] in 1987, originally referring to a parameter constructed by the weighted difference operation of relative variations in P-wave and S-wave velocities.[1] Goodway *et al.*[2] proposed the Lame's Modulus and Density technique, the core of which lies in using P-wave and S-wave velocity information from seismic data to invert for rock Lame parameters.[2] These parameters can provide more direct petrophysical property information than traditional P-wave and S-wave velocities, thereby enabling more effective identification of reservoir fluids and lithology. Patrick Connolly[3] proposed the concept of elastic impedance, which extended fluid identification research from the reflection coefficient domain to the impedance domain, significantly improving the accuracy and efficiency of reservoir fluid identification using seismic data.[3] In 2003, George Smith[4] proposed two fluid factor concepts: the fluid factor angle and the crossplot angle. Based on pre-stack amplitude versus offset (AVO) analysis, these concepts identify pore fluids by calculating angles related to fluid factors and crossplots.[4] To address the challenge that AVO responses in carbonate reservoirs are weaker than those in clastic reservoirs—making it difficult to accurately determine fluid properties—Peng *et al.*[5] proposed an innovative multi-angle elastic impedance fluid identification method.[5] Zheng *et al.*[6] constructed a new fluid factor formula based on fluid factors and Poisson impedance, which improves the sensitivity of fluid identification and lithology prediction.[6] Zong *et al.*[7] proposed a simultaneous inversion method named Fluid Factor, Mu (shear modulus), Rho (density)-Amplitude Variation with Angle based on partial angle stack seismic gather. This method is applicable for direct inversion of fluid factors, shear modulus, and density in heterogeneous reservoirs.[7] Yin *et al.*[8] summarized the key scientific issues in seismic fluid identification, focusing on reviewing the main progress of rock physics-driven seismic fluid identification, and discussed the related opportunities, challenges, and future research directions in seismic fluid identification.[8] Li *et al.*[9] derived a two-term approximate seismic reflection coefficient equation containing Gassmann fluid and shear modulus terms, based on Zoeppritz equation approximation and petrophysical relationships of elastic parameters. Under the Bayesian inversion framework, they developed a probabilistic pre-stack inversion method, using this equation, that excludes large-angle seismic data for direct inversion of the Gassmann fluid term and reservoir fluid identification[9] Jiang *et al.*[10] proposed to simplify the bulk modulus of saturated rocks using fluid bulk modulus and porosity, derived the linearized P-wave reflection coefficient as a function of fluid bulk modulus and modified porosity, and then established an inversion method to estimate fluid bulk modulus and porosity from observed seismic data using the obtained reflection coefficients and elastic impedance.[10] Liu *et al.*[11] proposed a new fluid factor by adding a correct term to the J attribute, and based on the neural network model, used deep learning methods to train and predict this new fluid factor.[11] In 2020, Li and Yin[12] proposed a mixed probability model-driven pre-stack seismic inversion method, providing a new technical approach for fluid factor extraction.[12] Farfour and Castagna *et al.*[13] introduced a new expression for fluid factors commonly used in AVO analysis and interpretation. The expression is a function of the common AVO intercept, gradient, and a weighting coefficient, which can be used to suppress the influence of lithology and other factors unrelated to fluid content.[13] Wu *et al.*[14] conducted joint inversion of fluid factor and brittleness index based on pre-stack seismic records from typical coal-measure gas blocks in the Sichuan Basin.[14] Lin *et al.*[15] first proposed a new method that treats the dynamic vertical reflectivity series as a dynamic variable varying with depth and lithology, and applied it to fluid factor calculation and inversion.[15] Zhang *et al.*[16] explored the concept of Q elastic impedance fluid factor and established a quantitative relationship between elastic impedance and Q elastic fluid factor for reservoir fluid identification.[16] Pan *et al.*[17] proposed an inversion method based on the elastic impedance of vertically transverse isotropic media, which can reasonably and accurately predict the fluid factor, Young's modulus, and anisotropic parameters, providing a new approach for fluid identification in shale reservoirs, prediction of brittleness parameters, and assessment of the development degree of anisotropy.[17]

Second, regularization inversion methods are crucial for improving the accuracy and reliability of subsurface fluid factor identification. They can suppress noise and instability by introducing constraints during the inversion process, enhancing the continuity and physical interpretability of inversion results. The concept of regularization was first proposed by Andrey Tikhonov[18] in 1943, and Tikhonov further developed the Tikhonov regularization method in 1963.[19] In 1992, Rudin *et al.* proposed the total variation regularization method for image denoising, particularly in scenarios affected by additive white Gaussian noise.[20] In 2005, Zou and Hastie[21] proposed the elastic net regularization, also known as L1+L2 regularization. Gholami[22] proposed an innovative nonlinear multichannel impedance inversion technique, which employs a total variation regularization strategy to guide the inversion process, promoting blocky features in the solutions. This method effectively highlights the boundaries of geological bodies and reduces random local inhomogeneities in data.[22] Ruixue *et al.*[23] proposed an innovative two-parameter shaping regularization method, which integrates the advantages of Tikhonov regularization and Total variation regularization. In 2016, Mousavi and Langston[24] presented a hybrid seismic denoising method in their research, combining higher-order statistics with an improved wavelet block thresholding technique. This method enhances denoising accuracy and efficiency while preserving effective signals in seismic data. Aiming at the issue that regularization constraints ignore the amplitude information of stratigraphic boundaries, Pan *et al.*[25] proposed an anisotropic total variation regularization with Lp-norm minimization (ATpV) pre-stack three-parameter inversion method based on reweighted L1. For the 1st time, this method integrates the reweighted L1 approach with the ATpV method and introduces it into pre-stack inversion, effectively restoring the sparsity of inversion results and improving inversion accuracy.[25]

This paper proposes a fluid factor inversion method based on frequency-domain two-step sub-band regularization, mainly embodied in the integration of regularization methods with multi-channel inversion theory in the frequency domain. It innovatively presents a frequency-domain lateral constraint regularization inversion method and further extends it. On this basis, aiming at the characteristics of seismic data, a two-step sub-band regularization method is proposed. The two-step sub-band regularization method first uses low-frequency band inversion to obtain the general morphology of subsurface structures, then performs high-frequency band inversion based on this model to acquire detailed geological details. Theoretical model verification shows that it processes low-frequency and high-frequency information more delicately, with significantly improved accuracy and noise resistance. Overall, compared with the traditional time–domain regularization multi-channel inversion method, it can provide more accurate inversion results during abrupt changes in subsurface media.

2. Methodology

2.1. The reflection coefficient equation incorporating the fluid factor

Wang *et al.*[26] clearly pointed out that the fluid factor, defined as the ratio of Lame coefficients to shear modulus (λ/μ)., exhibits higher sensitivity to fluid properties and lower sensitivity to porosity compared with other fluid factors. This characteristic enables effective differentiation between oil and water layers, enhancing the success rate of hydrocarbon detection. Therefore, this paper uses the ratio of Lame coefficients to shear modulus (λ/μ) as the fluid factor for inversion research in subsequent sections. Consequently, the GRAY formula involving these two parameters is introduced to derive the reflection coefficient equation incorporating the fluid factor. It should be noted that when facing geological bodies with high impedance contrast, the inversion accuracy of the GARY approximation will decrease with the increase of the incident angle. Therefore, the angle of multi-angle seismic data used for inversion should not be too large. The basic dimensions of seismic data are defined as follows: M is the number of incident angles, N is the number of seismic traces, and T is the number of sampling points per trace. The logarithmic vectors of the core parameters, namely the fluid factor λ/μ., shear modulus μ., and density ρ., are $L_{\lambda/\mu.}$, $L_{\mu.}$, and $L_{\rho.}$, respectively, (all are T×N matrices). All matrix dimensions involved in this paper are determined based on this basic dimensional framework.

The GRAY approximation is as follows in **Equation (I)**:

$$R(\theta)=(\frac{1}{4}-\frac{1}{2}k^2)sec^2\theta\frac{\Delta\lambda}{\lambda}+k^2(\frac{1}{2}sec^2\theta-2sin^2\theta)\frac{\Delta\mu}{\mu}+(\frac{1}{2}-\frac{1}{4}sec^2\theta)\frac{\Delta\rho}{\rho} \tag{I}$$

where, θ is the incident angle, in degrees; λ is the Lame coefficient; μ is the shear modulus, in Pa; ρ is the density, $k = V_s/V_p$ is the ratio of S-wave to P-wave velocity; $R_{\_\lambda} = \Delta\lambda/\lambda$ is the λ .reflection coefficient; $R_{\_\mu} = \Delta\mu/\mu$ is the μ.reflection coefficient; and $R_{\_\rho} = \Delta\rho/\rho$ is the density reflection coefficient.

Derivation with respect to λ/μ gives **Equation (II)**:

$$d(\lambda/\mu)=d(\lambda)\frac{\partial(\lambda/\mu)}{\partial\lambda}+d(\mu)\frac{\partial(\lambda/\mu)}{\partial\mu} \tag{II}$$

Dividing both sides of **Equation (II)** by λ/μ gives **Equations (III)** and **(IV)**:

$$\frac{d(\lambda/\mu)}{\lambda/\mu}=\frac{d(\lambda)\frac{\partial(\lambda/\mu)}{\partial\lambda}}{\lambda/\mu}+\frac{d(\mu)\frac{\partial(\lambda/\mu)}{\partial\mu}}{\lambda/\mu}=\frac{d(\lambda)}{\lambda}-\frac{d(\mu)}{\mu} \tag{III}$$

$$\frac{\Delta(\lambda/\mu)}{\lambda/\mu}=\frac{\Delta(\lambda)}{\lambda}-\frac{\Delta(\mu)}{\mu} \tag{IV}$$

Substituting **Equation (IV)** into **Equation (I)** and simplifying them yields **Equation (V)**:

$$R(\theta)=(\frac{1}{4}-\frac{1}{2}k^2)sec^2\theta\frac{\Delta(\lambda/\mu)}{\lambda/\mu}+(\frac{1}{4}sec^2\theta-2k^2sin^2\theta)$$
$$\frac{\Delta\mu}{\mu}+(\frac{1}{2}-\frac{1}{4}sec^2\theta)\frac{\Delta\rho}{\rho}$$
$$=a(\theta)R_{\_\lambda/\mu}+b(\theta)R_{\_\mu}+c(\theta)R_{\_\rho}$$
$$=a(\theta)R_{\_\lambda/\mu}+b(\theta)R_{\_\mu}+c(\theta)R_{\_\rho} \tag{V}$$

Where $R_{\_\lambda/\mu}$ represents the fluid factor λ/μ reflection coefficient. **Equation (V)** establishes a direct linear relationship among the λ/μ reflection coefficient, μ reflection coefficient, density reflection coefficient, and P-P reflection coefficient. Next, the objective function for the fluid factor λ/μ is constructed through **Equation (V)**. First, by establishing the relationship between seismic data and the parameters to be inverted through the reflection coefficient equation, we can obtain **Equation (VI)**:

$$S = AL \tag{VI}$$

Where S .is the multi-angle seismic data volume, L is the model parameter, and A is the forward operator.

2.2. The principle of the two-step sub-band regularization method

Based on the forward modeling of **Equation (VI)** derived from the above transformations, the traditional objective function is expressed as **Equation (VII)**:

$$J(L)=\min_L AL-S_2^2 \tag{VII}$$

Where $\|\cdot\|_2^2$ is the L2 norm. By incorporating the initial model L_0 into the objective function, the constraint of the initial model term is realized. Finally, by adding the time–domain L1 regularization method to the objective function, the objective function of the time–domain L1 regularization method can be expressed as **Equation (VIII)**:

$$J(L)=\min_L AL-S_2^2+\mu_{reg}L-L_{02}^{\,2}+\lambda_{reg}L_1 \tag{VIII}$$

Considering the strong correlation between adjacent seismic traces, the difference operator *D* is introduced. The time–domain horizontal constraint L1 regularization method adds an additional horizontal constraint term *D* based on **Equation VIII**. Thus, the objective function of the time-domain horizontal constraint L1 regularization inversion method is presented as **Equation (IX)**:

$$J(L)=\min_L AL-S_2^2+\mu_{reg}L-L_{02}^{\,2}+\lambda_{reg}DL_1 \tag{IX}$$

Where μ_{reg} is the regularization parameter for the initial model, the initial model L_0 is typically obtained by smoothing logging data, λ_{reg} is the regularization factor, and $\|\cdot\|_1$ is the L1 norm.

When abrupt changes occur in the stratum structure, the limitations of the time–domain horizontal constraint method become significantly prominent. Therefore, the inversion algorithm is extended to the frequency domain. Time domain signals are a combination of high and low frequencies, with low correlation between adjacent data. The residual of the differential operator *D* is susceptible to "frequency confusion errors," where horizontal constraints make it difficult to distinguish real geological changes from noise interference, leading to misjudgment of formation mutations and noise. After frequency band division, the correlation between data within the same frequency band is significantly enhanced, which can accurately distinguish geological mutations and interference. The frequency-domain regularization method achieves efficient conversion between the time–frequency domains by introducing the Fourier transform matrix *F* Substituting this matrix into the objective function, the objective function expression of the frequency-domain L1 horizontal constraint regularization inversion method is constructed as follows in **Equation (X)**:

$$J(L)=\min_L FAL-FS_2^2+\mu_{reg}FL-FL_{02}^{\,2}+\lambda_{reg}DFL_1 \tag{X}$$

Based on the difference in information contained in the low-frequency and high-frequency bands of seismic data, this paper innovatively proposes a two-step sub-band regularization inversion method within the framework of the frequency-domain L1 regularization inversion method. In the two-step method, the objective function is clearly divided into two independent parts, which respectively process the seismic data in the low-frequency and high-frequency bands. When processing data, the main frequency of seismic data is first obtained through spectral analysis. Then, the seismic data is divided into low-frequency and high-frequency parts by setting high-frequency and low-frequency limits. To avoid significant frequency truncation errors, modules and filters from mature commercial software can be directly used. First, an objective function with low-frequency seismic data is constructed, and the low-frequency band model L_{low}.is inverted by relying on the low-frequency seismic data S_{low}. (**Equation [XI]**):

$$J(L_{low})=\min_L FAL-FS_{low2}^{\;\;\;2}+\mu_{reg}FL-FL_{02}^{\,2}+\lambda_{reg}DFL_1 \tag{XI}$$

Finally, the obtained low-frequency band model L_{low}. is incorporated into the initial model constraint term to replace $L_{0.}$, and when combined with high-frequency seismic data, the objective function for the second stage can be obtained as **Equation (XII)**:

$$J(L)=\min_L FAL-FS_{high2}^{\;\;\;\;2}+\mu_{reg}FL-FL_{low2}^{\;\;\;2}+\lambda_{reg}DFL_1 \tag{XII}$$

Where F denotes the Fourier transform matrix, S_{low} denotes the seismic data in the low-frequency part, and S_{high} denotes the seismic data in the high-frequency part. The Fourier transform itself is a global transform, but combined with the proposed two-step sub-band regularization, it can achieve local constraints in the frequency domain, which is a mathematical advantage that cannot be achieved in the time domain.

Finally, to solve the two-step sub-band regularization inversion method, a matrix R_{low} is first introduced to make it equal to DFL .in **Equation (XII)**. Then, a quadratic penalty term and a dual term are added to transform **Equation (XII)** into an unconstrained optimization problem:

$$J(L_{low})=\min_{L} FAL-FS_{low}{}_2^2+\mu_{reg}FL-FL_0{}_2^2+\lambda_{reg}R_{low}{}_1$$
$$+\eta R_{low}-DFL+C_{low}{}_2^2 \qquad \text{(XIII)}$$

Similarly, using the same transformation method, **Equation (XIII)** can be converted into the following unconstrained problem in **Equation (XIV)**:

$$J(L)=\min_{L} FAL-FS_{high}{}_2^2+\mu_{reg}FL-FL_{low}{}_2^2+\lambda_{reg}R_{high}{}_1$$
$$+\eta R_{high}-DFL+C_{high}{}_2^2 \qquad \text{(XIV)}$$

It should be noted that the selection of parameters relies on multiple experiments to select empirical values. Among them are μ_{reg}=0.005–0.02, λ_{reg}=0.01–0.04, and η = 0.001–0.01. In the selection of the L1 regularization parameter λ_{reg}, a larger value should be set in the low-frequency range to strengthen constraints and ensure the overall continuity of the formation, while a smaller value should be set in the high-frequency range to weaken constraints and preserve the details of geological changes. An example of the impact of μ_{reg} and λ_{reg} on the inversion results can be found in the Appendix. Usually, within a given reference range, the inversion error increases with the increase of μ_{reg} and decreases with the rise in λ_{reg}. Among them, the influence of u on the inversion results is relatively small, less than the influence of r on the inversion results.

Then, using the alternating direction method of multipliers algorithm and combining **Equations (XIII)** and **(XIV)**, it is decomposed into the following subproblems (**Equations** [**XV**]–[**XX**]):

$$J(L_{low})=\min_{L} FAL-FS_{low}{}_2^2+\mu_{reg}FL-FL_0{}_2^2+\eta R_{low}$$
$$-DFL+C_{low}{}_2^2 \qquad \text{(XV)}$$

$$J(L)=\min_{L} FAL-FS_{high}{}_2^2+\mu_{reg}FL-FL_{low}{}_2^2+\eta R_{high}$$
$$-DFL+C_{high}{}_2^2 \qquad \text{(XVI)}$$

$$J(R_{low})=\lambda_{reg}R_{low}{}_1+\eta R_{low}-DFL+C_{low}{}_2^2 \qquad \text{(XVII)}$$

$$J(R_{high})=\lambda_{reg}R_{high}{}_1+\eta R_{high}-DFL+C_{high}{}_2^2 \qquad \text{(XVIII)}$$

$$J(C_{low})=\eta R_{low}-DFL+C_{low}{}_2^2 \qquad \text{(IX)}$$

$$J(C_{high})=\eta R_{high}-DFL+C_{high}{}_2^2 \qquad \text{(XX)}$$

Finally, the above subfunctions are solved. The workflow of the two-step sub-band regularization method in the frequency domain is shown in Figure 1.

3. Results

3.1. Synthetic data testing

To verify the inversion performance of the method, a fragment of the Marmousi2 model was selected to demonstrate the inversion effect, as shown in Figure 2. The model contained 5,000 traces horizontally, with 2,000 sampling points per trace. Based on the Marmousi2 theoretical model, reflection coefficients at three different incident angles (10°, 20°, and 30°) were calculated, which were then convolved with a zero-phase wavelet with a dominant frequency of 30 Hz to synthesize multi-angle seismic data volumes. Figure 3 shows the angular seismic data volumes corresponding to these three incident angles (10°, 20°, and 30°), respectively.

Next, the traditional time–domain horizontal constraint L1 regularization method and the proposed two-step sub-band regularization method were applied to a two-dimensional model. As shown in Figure 4, the inversion results of the ratio of Lame coefficient to shear modulus obtained by the time-domain L1 regularization and time–domain horizontal constraint L1 regularization methods are displayed.

To more intuitively compare the inversion effects of the two methods, as shown in Figure 5, the inversion errors for each method were calculated by measuring the differences between the original model and the results of the two inversion methods.

From the images in the red box of Figure 5, it can be seen that the area of bright regions in Figure 5B is smaller than that in Figure 5A, and the light blue line to the left of the red box in Figure 5B is thinner than the line at the same position in Figure 5A. This indicates that the error of the two-step sub-band regularization method is lower than that of the time–domain L1 horizontal constraint regularization method, suggesting that the two-step sub-band regularization method outperforms the time–domain L1 horizontal constraint regularization method in inverting fluid factors. This difference is

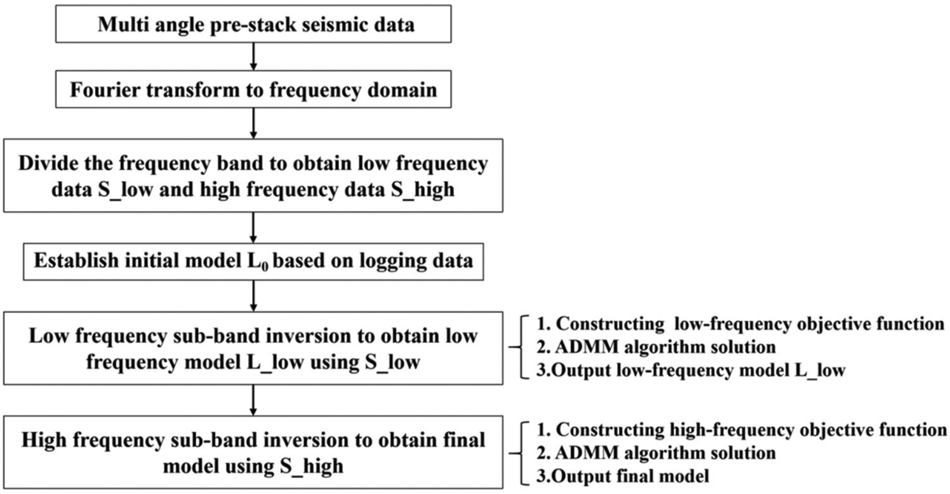


Figure 1. Workflow diagram of the two-step sub-band regularization inversion
Abbreviation: ADMM: Alternating direction method of multipliers.

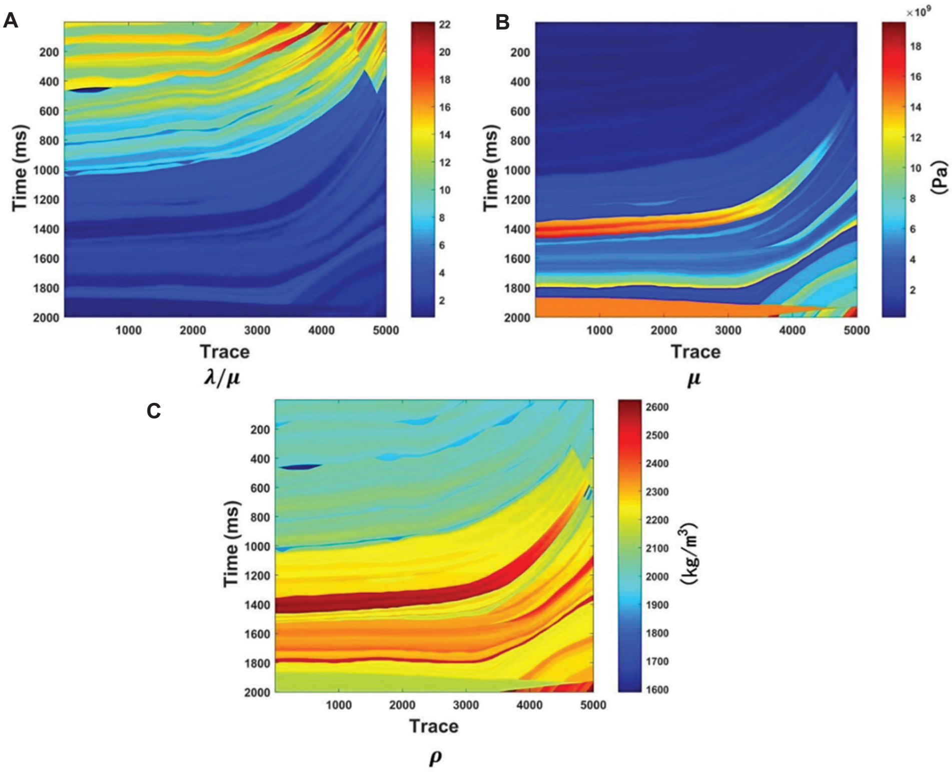


Figure 2. Theoretical model. (A) Ratio of Lame coefficient to shear modulus; (B) shear modulus; and (C) density,

attributed to the more refined sub-band processing and optimization of the two-step method, which more effectively reduces errors and improves inversion accuracy.

Furthermore, Figures 6 and 7 also show the error images of the inversion results for shear modulus and density using the two regularization methods. In terms of the overall inversion error, the sub-band regularization method exhibited lower errors than the time-domain L1 horizontal constraint regularization method, demonstrating the effectiveness of the sub-band regularization method.

To demonstrate the effects and differences between the time–domain L1 horizontal constraint regularization

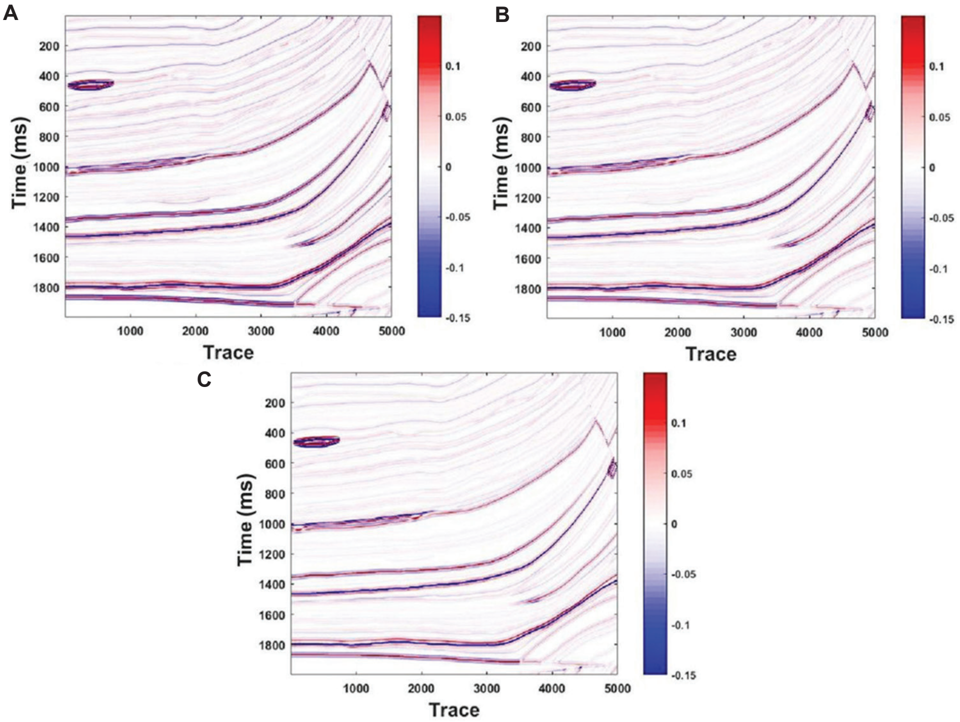


Figure 3. Synthetic seismic records at different incident angles. (A) 10°; (B) 20°; and (C) 30°.

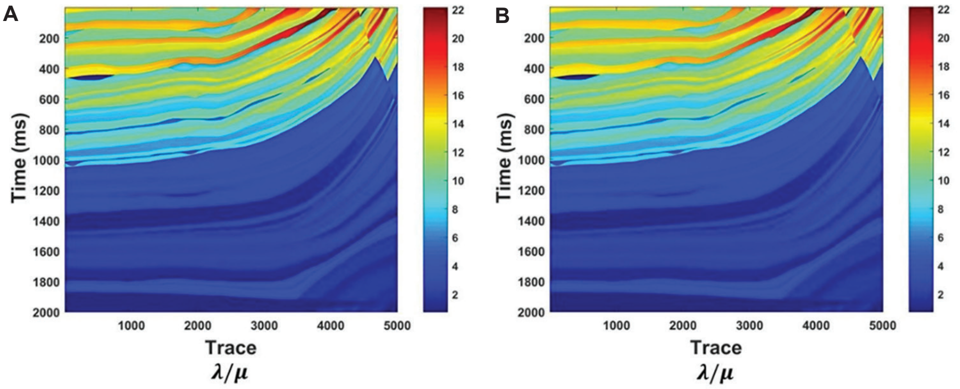


Figure 4. λ/μ inversion results. (A) The time–domain L1 horizontal constraint regularization; and (B) the two-step sub-band regularization.

method and the sub-band regularization method in more detail, this paper extracts the 1,000th trace data of the used model for single-trace analysis, as shown in Figure 8. From the single-trace inversion effect, it can be clearly seen that the single-trace inversion record obtained by the sub-band regularization method has a higher matching degree with the original data, more minor inversion error, and higher accuracy. To more intuitively demonstrate the difference in inversion performance between the time–domain L1 horizontal constraint regularization method and the two-step sub-band regularization method. Table 1 details the root mean square error (RMSE) of the inversion results obtained by these two methods.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(L - L_{model})^2}{n}} \quad \text{(XXI)}$$

In **Equation (XXI)**, L represents the inversion result and L_{model} represents the theoretical model of the inversion. The smaller the RMSE, the higher the inversion accuracy of this method. By comparing the data in the table, the two-step sub-band regularization method has lower RMSE in the inversion results of the three parameters compared to the

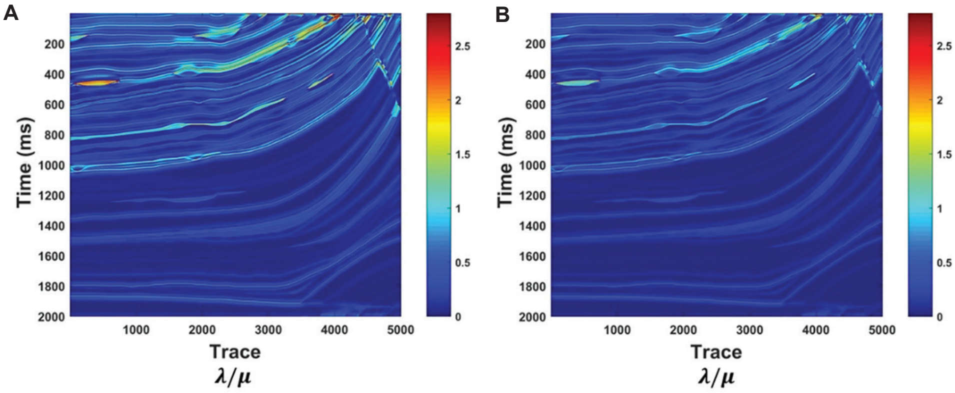


Figure 5. λ/μ inversion error. (A) The time-domain L1 horizontal constraint regularization; and (B) the two-step sub-band regularization.

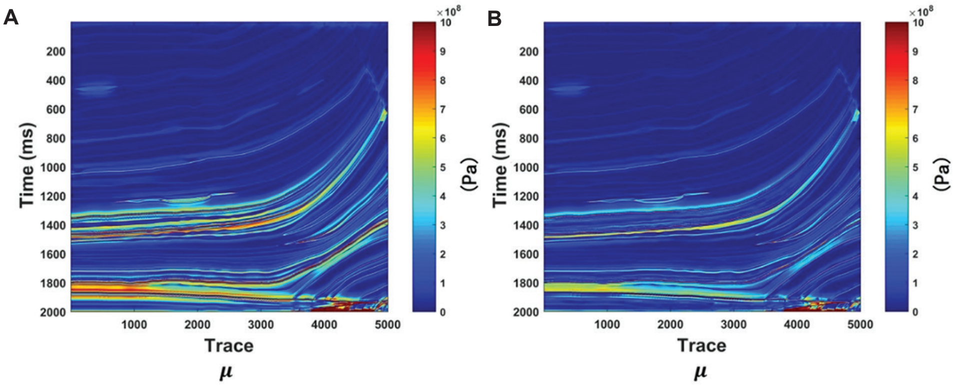


Figure 6. μ inversion error. (A) The time–domain L1 horizontal constraint regularization; and (B) the two-step sub-band regularization.

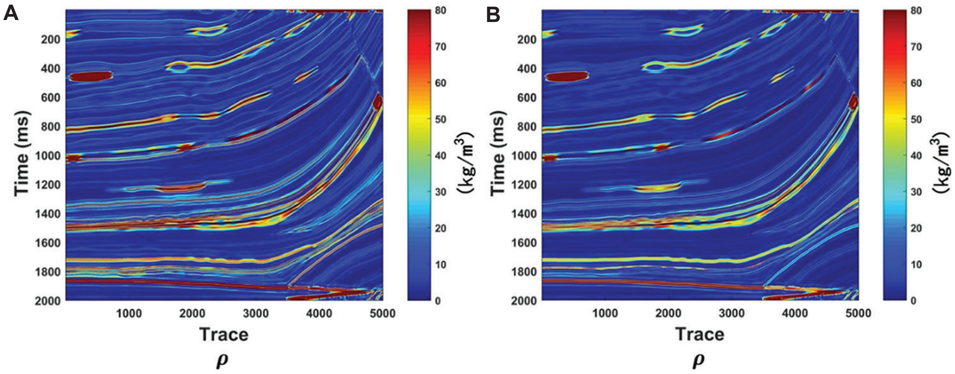


Figure 7. ρ inversion error. (A) The time-domain L1 horizontal constraint regularization; and (B) the two-step sub-band regularization.

time–domain L1 horizontal constraint method. This result fully indicates that the sub-band regularization method is superior to the time–domain L1 horizontal constraint regularization method in terms of inversion effect and can more effectively improve the inversion accuracy.

In addition, we also compared the running time of the two methods. Due to the larger computational complexity and the need for two inversions, the frequency domain method has a higher time consumption, as clearly reflected in Table 2. After multiple experimental verifications, the time consumption of the proposed method is about twice that of the L1 method in the time domain. The entire experimental paper is based on a computer processing unit, Intel (R) Xeon (R) CPU E5-2680 v4 × 2 @ 2.40GHz.

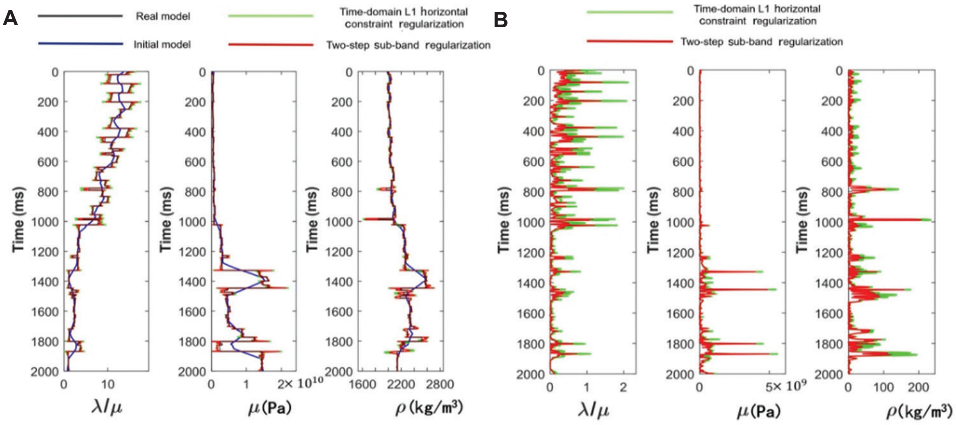


Figure 8. Single-trace data inversion results and errors. (A) Inversion results and (B) inversion result errors.

Table 1. Root mean square error comparison of 5,000 traces inversion results from different methods

Method	Root mean square error		
	λ/μ	μ (*Pa*)	ρ (*kg/m³*)
Time-domain L1 horizontal constraint regularization	0.013	1.0687×10^7	1.0123
Two-step sub-band regularization	0.0064	0.7788×10^7	0.6096

Table 2. Comparison of the time consumption of different methods

Method	Time	Data size
Time-domain L1 horizontal constraint regularization	12 min	5,000×2,000
Two-step sub-band regularization	26 min	5,000×2,000

To further verify the noise resistance performance of the method, a noise tolerance experiment was conducted with an additional 5dB signal-to-noise ratio (SNR) of Gaussian white noise, considering that field seismic data contain certain noise. Figure 9 shows the comparison of inversion results from different methods after the addition of noise. The analysis indicates that when the SNR is 5dB, abnormal features appeared in the inversion results, and the inversion errors of both regularization methods exhibited an increase in abnormal points. However, compared with the time–domain L1 regularization method, the proposed two-step sub-band regularization method demonstrated superior performance in terms of continuity and stability of inversion results, with more minor errors and a more uniform distribution. The experimental results confirm that the two-step sub-band regularization method can still maintain high inversion accuracy and reliability under low SNR conditions, fully highlighting the advantages of this method. It should be noted that the current method is only effective on random noise. If there is coherent noise in the data, denoising methods should be applied first for noise suppression. This is also consistent with the situation in field data inversion (where the noise after data denoising is mostly random noise).

3.2. Field data testing

To further verify the effectiveness of the method, the time–domain horizontal constraint L1 regularization method and the two-step sub-band regularization method were applied to an actual work area. These field data were from a shale gas development area in the Sichuan Basin, China. We presented some well-crossing angle stack sections. Figure 10 shows the seismic data volumes with incident angles of 12°, 24°, and 36°, respectively. The section contained 793 traces with a trace interval of 20 m and a vertical time sampling interval of 2 ms.

In the process of seismic inversion, smoothing of logging data is a key step in constructing the initial model. First, the initial model of the work area can be obtained by smoothing the logging data, as shown in Figure 11. This model provides an essential foundation for subsequent seismic data inversion.

As shown in Figures 12-14, the inversion results of fluid factors, shear modulus, and density are displayed respectively. Figures 12-14A show the results of the time–domain horizontal constraint L1 regularization method, and Figures 12-14B show those of the two-step sub-band regularization method. From the inversion results, it can be seen that compared with the time–domain horizontal constraint L1 regularization method, the two-step sub-band regularization method significantly improved the lateral continuity of inversion results. Meanwhile, the vertical resolution is effectively enhanced, and the overall performance is notably improved.

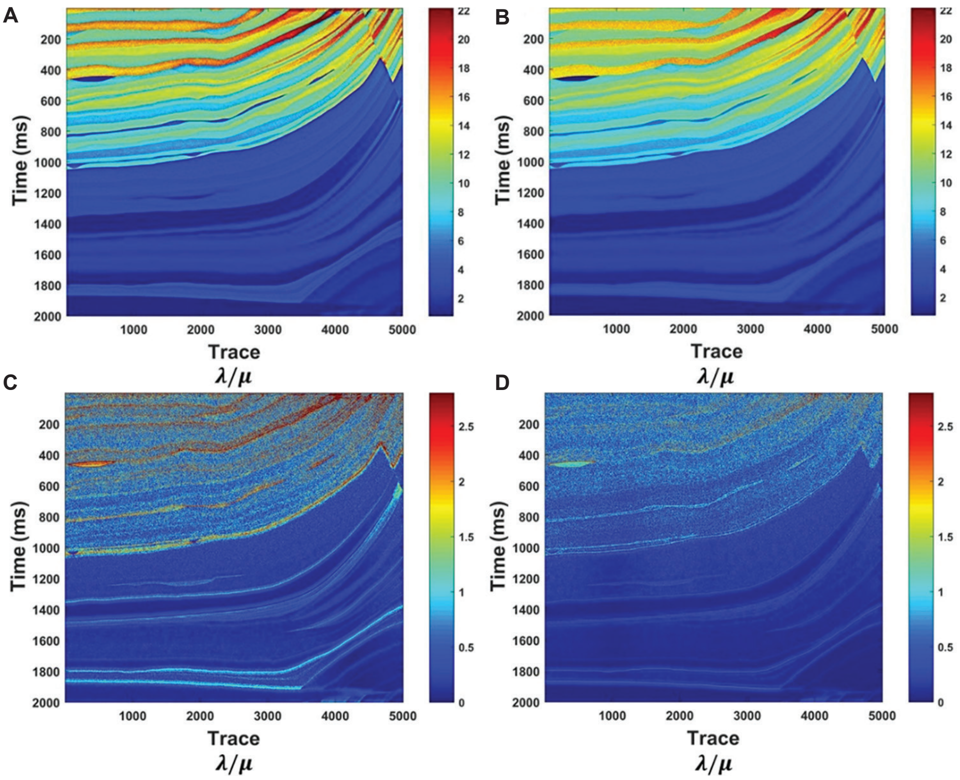


Figure 9. Inversion results and errors of fluid factors with added noise. (A) Inversion results of time–domain L1 horizontal constraint regularization; (B) inversion results of two-step sub-band regularization; (C) inversion errors of time–domain L1 horizontal constraint regularization; and (D) inversion errors of two-step sub-band regularization.

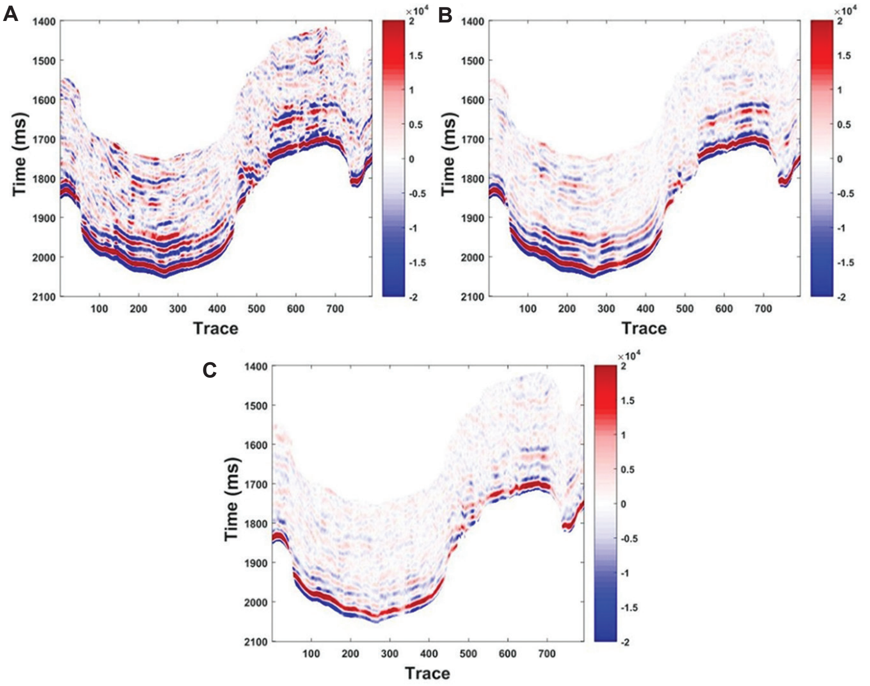


Figure 10. Angle stack section of field data. (A) 12°; (B) 24°; and (C) 36°.

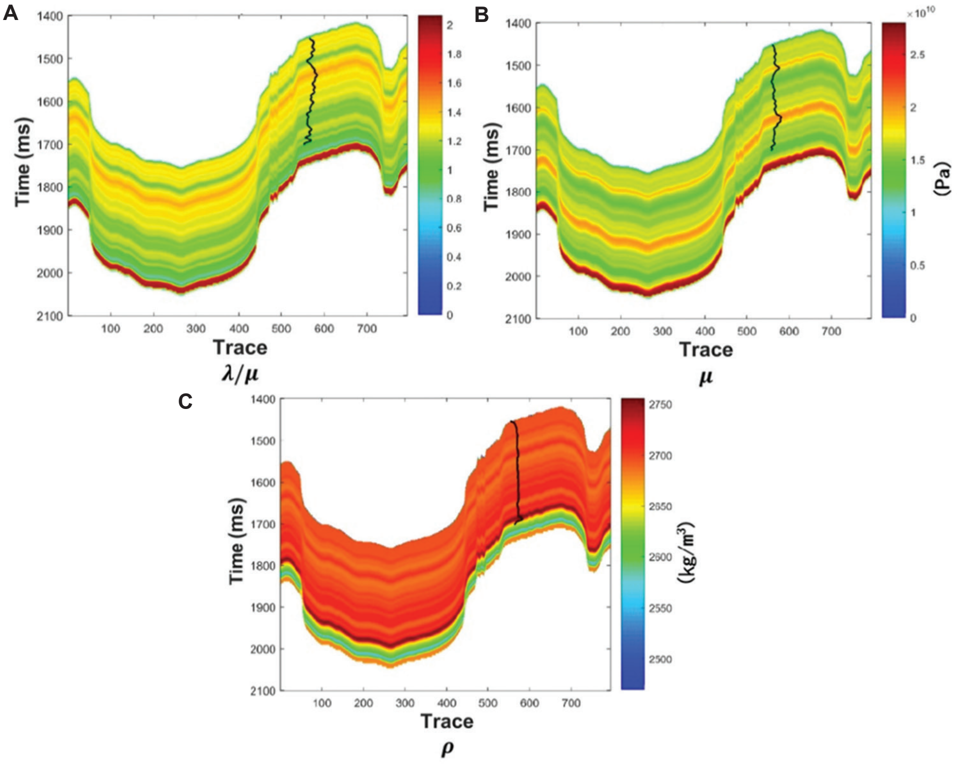


Figure 11. Initial model. (A) λ/μ.; (B) μ.; and (C) ρ.

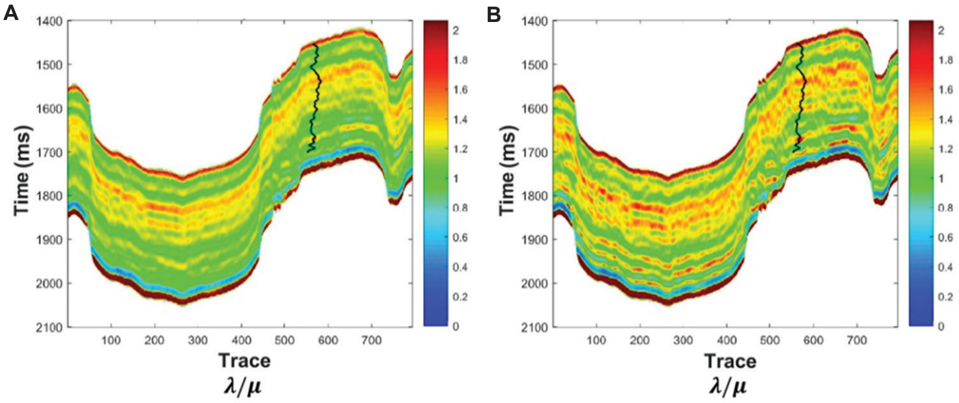


Figure 12. λ/μ inversion results. (A) The time–domain L1 horizontal constraint regularization; and (B) the two-step sub-band regularization.

To verify the reliability of the inversion results, we extracted the well-side traces and conducted a detailed comparison with the logging data. Figure 15 presents the inversion comparison results of the two methods: The black curve represents the actual logging curve, the blue curve denotes the inversion result of the time–domain horizontal constraint L1 regularization method, and the red curve stands for the inversion result of the two-step sub-band regularization method. For a more detailed comparison, Table 3 shows the correlation coefficients between the inversion results and the logging data. The results show that although the two-step sub-band regularization method achieved a limited improvement, its inversion results were closer to the logging data and exhibited higher stability. This indicates that the method can more accurately capture geological features during the inversion process, thereby providing more reliable and precise inversion results.

Regarding the relatively slight improvement effect shown in Figure 15, we speculate that it mainly stems from two reasons: first, as observed from the original seismic data, the resolution of the well-side data itself

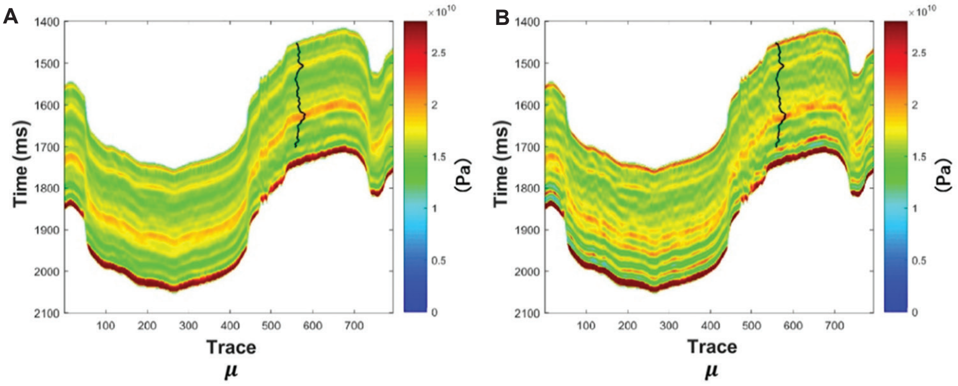


Figure 13. μ inversion results. (A) The time–domain L1 horizontal constraint regularization; and (B) the two-step sub-band regularization.

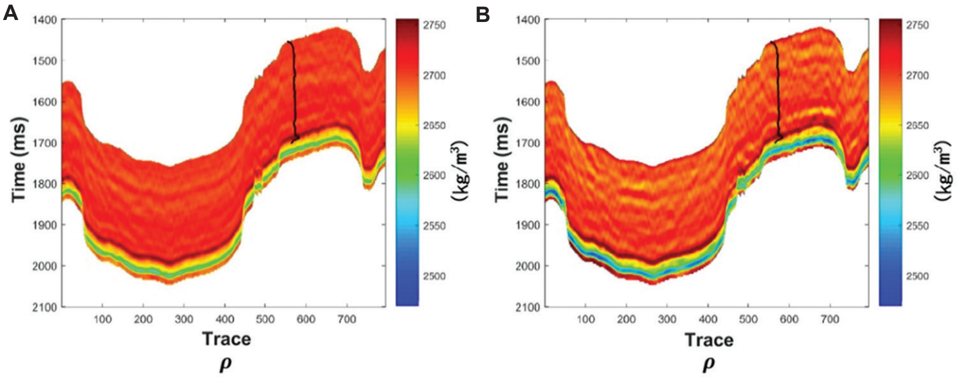


Figure 14. ρ inversion results. (A) The time–domain L1 horizontal constraint regularization; and (B) the two-step sub-band regularization.

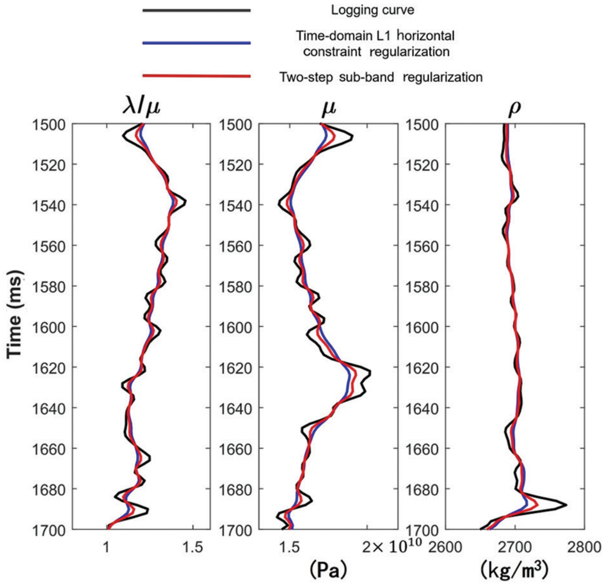


Figure 15. Comparison of well-side trace inversion results

has limitations, which to a certain extent restricts the full exertion of the performance of the proposed method; second, there is a deviation between the extracted wavelet and the real wavelet, and this factor also affects the actual inversion effect.

Table 3. Comparison of correlation coefficients between inversion results and logging data

Method	Root mean square error		
	λ/μ	μ	ρ
Time–domain L1 horizontal constraint regularization	0.6824	0.5973	0.7158
Two-step sub-band regularization	0.7456	0.6608	0.7326

4. Discussion

In response to the differences in information contained in different frequency bands of seismic data (low frequencies reflect structural outlines, while high frequencies depict detailed features), this paper groundbreakingly proposes a two-step frequency-band division regularization method. Through frequency-band division processing, geological information in different frequency bands was explicitly enhanced during inversion, avoiding information confusion caused by the unified processing of the whole frequency band. Meanwhile, the Alternating Direction Method of Multipliers was adopted to efficiently decompose and solve the objective function, ensuring the algorithm's convergence and computational efficiency. The results of the theoretical model experiments clearly verify the significant improvement of this method in inversion accuracy and noise resistance. This fully demonstrates the scientific validity and feasibility of the frequency-band division strategy, providing a new technical approach for the accurate inversion of fluid factors.

It should be further clarified that the proposed method has certain applicable conditions and limitations. First, the data needs to meet the requirements of multi-angle pre-stack seismic data and complete logging data, ensuring the reliability of frequency domain feature extraction and initial model. The second is that the geological structure adapts to sudden changes in the lateral strata without severe fragmentation. If there are large-scale thrust structures or global fractured zones, it will destroy the frequency domain data correlation and lead to the failure of lateral constraints. The third is that fluid differences can be distinguished by λ/μ. If the fluid differences are weak or the reservoir exceeds the range of seismic frequency band identification, the advantages of the method are difficult to reflect.

Furthermore, as a core factor of geological complexity, the impact of fracture development on inversion results requires focused attention, with specific variations depending on the degree of fracture development: In areas with low-to-moderate fracture development, microfractures only slightly alter elastic parameters and the overall λ/μ remains dominated by fluid type; the proposed method can capture weak anomalies of gas-bearing fractures through high-frequency inversion. In areas with high fracture development, seismic wave scattering blurs the low-high frequency stratification characteristics in the frequency domain, and fracture fillings may obscure the fluid indication significance of λ/μ, thereby affecting inversion results. In areas with extreme fracture development, where seismic wave propagation is dominated by scattering-diffraction, the linear relationship of the forward model fails, and inversion results cannot distinguish the contributions of fluids, fractures, and karst caves, leading to the fundamental failure of the method.

5. Conclusion

This study focused on pre-stack seismic inversion methods and proposes a new fluid factor inversion method based on a two-step sub-band regularization in the frequency domain. Starting from the ill-posed problem of pre-stack seismic inversion based on the constructed fluid factor reflection coefficient formula, the study successively constructs the time–domain horizontal constraint L1 regularization and frequency-domain horizontal constraint L1 regularization methods, finally developing the two-step sub-band regularization inversion method. Theoretical model and field data verification results showed that the new method has significant advantages in vertical resolution, lateral continuity, and noise resistance, outperforming the traditional time–domain horizontal constraint L1 regularization method. By directly inverting fluid factors, this method effectively reduces cumulative errors. Combined with the frequency-domain horizontal constraint mechanism, it minimizes inversion uncertainties. Experimental verification confirms its feasibility, providing a high-precision and robust inversion tool for fluid prediction under complex geological conditions.

Finally, it is worth noting that this study still has certain limitations, which need to be further improved in subsequent work. First, the selection of current regularization parameters relies on manual experience and multiple sets of experimental trials. Although it is operable in small-scale data experiments, it will require a significant amount of time and effort when applied to actual large-scale seismic data. In subsequent research, automated methods (such as L-curve analysis or cross-validation methods) can be considered to optimize the selection of regularization parameters. Secondly, the current process of dividing high and low frequencies is very subjective. In future research, it is necessary to develop algorithms to automatically divide high and low frequencies and discuss the impact of different frequency band divisions on the inversion results.

Acknowledgments

None.

Funding

This work was supported in part by the Technology Cooperation Project of the CNPC-SWPU Innovation Alliance (2020C×020000) and in part by the Natural Science Foundation of Sichuan Province (24NSFSC0808).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Peng Xiao
Formal analysis: Pang Chen
Investigation: Peng Xiao, Wangyang Xu
Methodology: Peng Xiao
Writing–original draft: Peng Zhang
Writing–review & editing: Ying Xiao

Availability of data

The data supporting the findings of this study are available from the corresponding author on reasonable request.

References

1. Smith GC, Gidlow PM. Weighted stacking for rock property estimation and detection of gas. *Geophys Prospect*. 1987;35(9):993-1014.

   doi: 10.1111/j.1365-2478.1987.tb00856.x
2. Goodway B, Chen T, Downton J. Improved AVO fluid detection and lithology discrimination using lame petrophysical parameters; "λρ, μρ, λμ fluid stack", From P and S inversions. *SEG Ann Meet*. 1997;22:183-186.

   doi: 10.1190/1.1885795
3. Connolly P. Elastic impedance. *Lead Edge*. 1999; 18(4):438-438.

   doi: 10.1190/1.1438307
4. Smith GC. *The Fluid Factor Angle and the Crossplot Angle*. *SEG Annual Meeting*. United States: Society of Exploration Geophysicists; 2003.

   doi: 10.1190/1.1817679
5. Peng ZM, Li YL, Wu SH, He ZH, Zhou YJ. Discriminating gas and water using multi-angle extended elastic impedance inversion in carbonate reservoirs. *Chin J Geophys*. 2008;51(3):881-885.

   doi: 10.1002/cjg2.1253
6. Zheng JJ, Yin XY, Zhang GZ. Fluid factor analysis and the construction of the new fluid factor. *Prog Geophys*. 2011;26(2):579-587.

   doi: 10.3969/j.issn.1004-2903.2011.02.024
7. Zong Z, Yin X, Wu G. Direct inversion for a fluid factor and its application in heterogeneous reservoirs. *Geophys Prospect*. 2013;61(5):998-1005.

   doi: 10.1111/1365-2478.12038
8. Yin X, Zong Z, Wu G. Research on seismic fluid identification driven by rock physics. *Sci China Earth Sci*. 2015;58(2):159-171.

   doi: 10.1007/s11430-014-4992-3
9. Li C, Zhang JM, Zhu ZY. Direct inversion method for fluid factor in deep reservoirs. *Petrol Geophys Prospect*. 2017;56(6):827-834.

   doi: 10.3969/j.issn.1000-1441.2017.06.008
10. Jiang Z, Xiong Y. Seismic inversion for fluid bulk modulus based on elastic impedance. *J Appl Geophys*. 2019;169:74-84.

    doi: 10.1016/j.jappgeo.2019.06.013
11. Liu C, Ghosh DP, Salim AMA, Chow WS. A new fluid factor and its application using a deep learning approach. *Geophys Prospect*. 2019;67(1):140-149.

    doi: 10.1111/1365-2478.12712
12. Li K, Yin XY, Zong ZY. Facies-constrained prestack seismic probabilistic inversion driven by rock physics. *Sci China Earth Sci*. 2020;50(6):832-854.

    doi: 10.1007/s11430-019-9578-1
13. Farfour M, Castagna JP. A new expression for fluid factor using AVO intercept and gradient. *Soc Explor Geophys*. 2021;9(1):267-271.

    doi: 10.1190/segam2021-3575102.1
14. Wu HB, Wu RX, Zhang PS, Huang Y, Dong S. Combined fluid factor and brittleness index inversion for coal-measure gas reservoies. *Geophys Prospect*. 2022;70(4):751-764.

    doi: 10.1111/1365-2478.13172
15. Zhou L, Liao JP, Liu XY, Wang P, Guo YN, Li JY. A high resolution inversion method for fluid factor with dynamic dry-rock VP/VS ratio squared. *Petrol Sci*. 2023;20(5):2822-2834.

    doi: 10.1016/j.petsci.2023.09.015
16. Ruyi Z, Huan W, Yongqiang M, Jianhua T, Chao H, Maoqiang Z. Fluid factor inversion based on *Q* elastic impedance. *Geophys Prospect Petrol*. 2024;63(4):826-832.

    doi: 10.12431/issn.1000-1441.2024.63.04.011
17. Pan F, Li SJ, Qin DW, *et al*. Direct inversion method for fluid factor and anisotropic parameters in VTI media. *Oil Geophys Prospect*. 2024;59(4):875-886.

    doi: 10.13810/j.cnki.issn.1000-7210.2024.04.025
18. Tikhonov AN. On the stability of inverse problems. *Dokl Akad Nauk Sssr*. 1943;39(5):195-198.
19. Tikhonov AN. Solution of incorrectly formulated problems and the regularization method. *Sov Math Dok*. 1963;4:1035-1038.
20. Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithm. *Phys D Non Phenom*. 1992;60(4):259-268.

    doi: 10.1016/0167-2789(92)90242-F
21. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Roy Statis Soc*. 2005;67(2):301-301.

    doi: 10.1111/j.1467-9868.2005.00503.x
22. Gholami A. Nonlinear multichannel impedance inversion by total-variation regularization. *Geophysics*. 2015;80(5):217-224.

    doi: 10.1190/geo2015-0004.1
23. Ruixue S, Liguo H, Eryan S, *et al*. Dual-parameter shaping regularized full waveform inversion in frequency domain. *Glob Geol*. 2015;18(4):258-262.

    doi: 10.3969/j.issn.1673-9736.2015.04.09

24. Mousavi SM, Langston CA, Horton SP. Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform. *Geophysics*. 2016;81(4):341-355.

    doi: 10.1190/geo2015-0598.1

25. Pan S, Chen Y, Yin C, Gou Q, Zhang D. Prestack inversion method with ATpV regularization based on reweighted L1. *J S Petrol Univ Sci Technol*. 2024;46(3):13-26.

    doi: 10.11885/j.issn.1674-5086.2022.08.20.02

26. Wang D, Zhang YM, Niu C, *et al*. The optimization of sensitive fluid factor removing the effect of porosity and its application to hydrocarbon detection. *Geophys Geochem Explor*. 2021;45(6):1402-1408.

    doi: 10.11720/wtyht.2021.1364

Appendix

A. Forward equation derivation

To accurately extract key parameters from seismic data, it is necessary to establish the connection between the seismic data and the parameters to be inverted using the reflection coefficient equation. Taking density ρ as an example, the reflection coefficient of density ρ in seismic data can be approximately expressed as **Equation (A1)**:

$$R_{\_\rho}=\frac{\rho_{i+1}-\rho_i}{\rho_{i+1}+\rho_i}\approx\frac{1}{2}\Delta ln(\rho)=\frac{1}{2}[ln(\rho_{i+1})-ln(\rho_i)] \tag{A1}$$

Let L_i be the natural logarithm of density ρ, and $L_i=ln(\rho)$ **(Equation [A2]**):

$$R_{\_\rho}\approx\frac{1}{2}(L_{i+1}-L_i) \tag{A2}$$

And $R_{\_\rho}$ can be expressed in matrix form **(Equation [A3]**):

$$R_{\_\rho}=\frac{1}{2}DL \tag{A3}$$

Where L is the natural logarithm of density and D is the difference matrix.

The construction of the seismic inversion objective function is based on the mathematical foundation of the convolution model. In this process, the wavelet convolution kernel matrix is first introduced to transform the wavelet convolution operation in seismic data into linear algebraic operations **(Equation [A4]**):

$$S=W\cdot R \tag{A4}$$

By using the angle-dependent wavelet matrix W(θ), the angle-dependent forward seismic record S(θ) can be expressed as **Equation [A5]**:

$$\begin{bmatrix} S(\theta) \\ \vdots \\ S(\theta_M) \end{bmatrix}=\frac{1}{2}\begin{bmatrix} W(\theta_1) & & \\ & \ddots & \\ & & W(\theta_M) \end{bmatrix}\begin{bmatrix} a(\theta_1) & b(\theta_1) & c(\theta_1) \\ \vdots & \vdots & \vdots \\ a(\theta_M) & b(\theta_M) & c(\theta_M) \end{bmatrix}\begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & -1 & 1 \end{bmatrix}\begin{bmatrix} L_{\lambda/\mu} \\ L_\mu \\ L_\rho \end{bmatrix} \tag{A5}$$

Where a(θ$_i$), b(θ$_i$), and c(θ$_i$) are all diagonal matrices, and the matrix composed of these three diagonal matrices is defined as *B*. Then, the forward modeling equation based on the linear approximation formula can be abbreviated as **Equation [A6]**:

$$S = WBDL = AL \tag{A6}$$

Where *A = WBD.*

B. An example of a sensitivity analysis comparison figure for regularization parameters

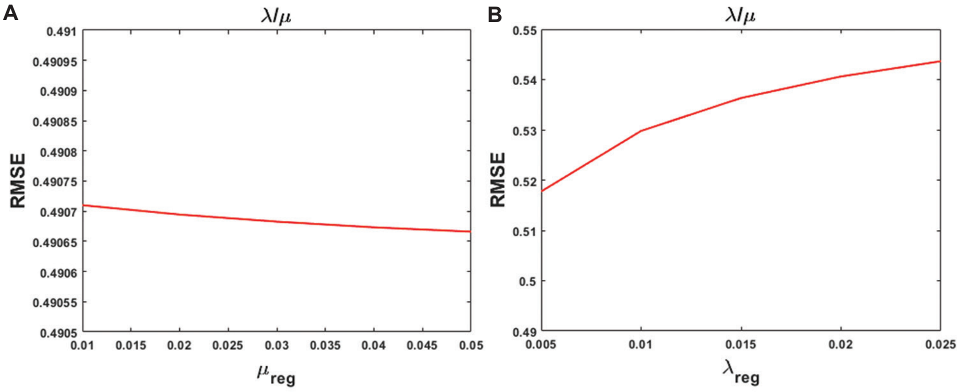


Figure A1. λ/μ inversion results. (A) The impact of μ_{reg} on the inversion results of λ/μ. (B) The impact of r_{reg} on the inversion results of λ/μ.
Abbreviation: RMSE: Root mean square error.

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Application of the spatial autocorrelation method incorporating damping optimization for identifying goafs in cultural heritage sites

Jiale Liu[1], **Xiaodong Wang**[1*], **Yanhai Liu**[2], **Ke Ren**[3], **and Shangqing Zhang**[4]

[1]Department of Geophysics and Information, College of Geoscience and Surveying Engineering, China University of Mining and Technology-Beijing, Beijing, China

[2]China Energy Shendong Coal Group Co., Ltd., Ordos, Nei Mongol Autonomous Region, China

[3]School of Remote Sensing Science and Technology, Aerospace Information Technology University, Jinan, Shandong, China

[4]Shanxi Key Laboratory of Metallogeny and Assessment of Strategic Mineral Resources, Taiyuan, Shanxi, China

Abstract

As a World Cultural Heritage Site, the Yungang Grottoes face the risk of geological disasters caused by underground goafs. To accurately detect the surrounding void zones of the grottoes and analyze their distribution and structural characteristics, this study uniquely applied the passive-source surface wave spatial autocorrelation (SPAC) method. Four high-density linear arrays were deployed in the Yungang Grottoes area to collect microtremor signals. By extracting the Rayleigh wave dispersion curve and combining it with a non-linear least-squares inversion technique incorporating a damping regularization term, a shallow shear wave velocity profile was established, successfully identifying potential low-velocity anomalies. During the inversion process, the optimized damping factor effectively suppressed the oscillation effects under complex geological conditions, significantly improving the stability and accuracy of the imaging. The results showed that 15 typical low-velocity anomalies with wave velocities below 1,200 m/s were identified, and these anomalies were highly consistent with the locations of goaf areas in historical mining data. The imaging results revealed that the goaf exhibits multiple layers and a multi-center distribution, with the L1 and L4 survey lines being the main areas of goaf activity. With a horizontal resolution of 10–20 m and a maximum detection depth of 320 m, the method presented differences in the integrity and fragmentation of underground rock masses. The SPAC method demonstrated advantages such as high resolution, non-destructive testing capability, and imaging stability in the detection of abandoned mine areas within cultural heritage sites. By optimizing the inversion regularization parameters, this study significantly improves imaging accuracy in complex geological environments, providing effective technical support for the stability assessment of the Yungang Grottoes and for the prevention and control of geological disaster risks at cultural heritage sites.

Keywords: Spatial autocorrelation method; Damping factor; Yungang Grottoes; Goaf

***Corresponding author:**
Xiaodong Wang
(bqt2300202017@student.cumtb.edu.cn)

Citation: Liu J, Wang X, Liu Y, Ren K, Zhang S. Application of the spatial autocorrelation method incorporating damping optimization for identifying goafs in cultural heritage sites. *J Seismic Explor*. 2025;34(5):18-35.
doi: 10.36922/JSE025330059

Received: August 16, 2025

Revised: October 16, 2025

Accepted: October 21, 2025

1. Introduction

As a large Buddhist cave complex built in the fifth century, the Yungang Grottoes are among the first sites to be included in the United Nations Educational, Scientific, and Cultural Organization (UNESCO) World Heritage List and are of outstanding historical, cultural, and artistic value. In recent years, due to the long-term mining of surrounding coal resources, the underground geological environment of the site has undergone significant changes.[1] In particular, the existence of goafs has become a major source of geological hazards, threatening the structural safety and long-term stability of the grottoes.[2] Accurate detection of the distribution range and structural characteristics of goafs is a key technical step in the effective protection of cultural relic sites and in the prevention and control of geological disaster risks.[3-6]

Globally, cultural heritage sites face increasingly severe threats from mining-induced geological hazards. The Wieliczka Salt Mine in Poland, a UNESCO World Heritage Site, has experienced terrain deformation due to centuries of salt extraction, with interferometric synthetic aperture radar (InSAR) technology employed for continuous monitoring.[7] In China, microtremor survey methods have been successfully applied to map collapsed columns and goafs in coal mining areas.[8,9] Despite these advances, achieving high-resolution subsurface imaging under complex geological conditions while maintaining non-invasive protocols remains challenging, particularly for heritage sites with strict construction constraints.

Various geophysical techniques have been developed for goaf detection. Martínez-Moreno *et al.*[10] used a combination of microgravity, two-dimensional (2D) electrical resistivity imaging (ERT), and induced polarization to reveal cavity structures in Spain's Algaidilla Cave. Metwaly and AlFouzan[11] applied 2D ERT for shallow goaf imaging in Saudi Arabia. Styles *et al.*[12] evaluated the applicability of microgravity in karst and abandoned mine detection. In cultural heritage contexts, passive-source methods offer particular advantages due to their non-invasive nature, strong anti-interference capability, and flexible deployment.[13-15] Zeid *et al.*[16] successfully applied the horizontal to vertical spectral ratio technology to identify ancient soil layer interfaces at Italy's Pilastri site, while Schwellenbach *et al.*[17] used passive surface waves at Sicily's Selinunte Archaeological Park to obtain high-resolution shallow velocity structures. However, inversion instability and velocity oscillations under abrupt gradient changes remain unresolved, limiting imaging quality and geological interpretation accuracy in complex scenarios.

To address these challenges, this paper introduces the spatial autocorrelation (SPAC) passive surface wave method for high-resolution goaf imaging in the Yungang Grottoes Protection Area. By extracting Rayleigh wave dispersion curves from continuous microtremor data acquired through multiple observation lines and employing multi-parameter inversion with damped regularization control, we establish shear wave velocity profiles to identify low-velocity anomaly zones indicative of goaf distributions. The adjustable damping factor suppresses non-geological oscillations caused by abrupt velocity changes, significantly improving imaging stability and geological interpretation reliability for complex structures. The research results not only fill the gap in knowledge about the near-surface structure of the region but also provide a scientific basis for assessing the potential impact of goafs on the stability of the Yungang Grottoes, offering technical support for long-term cultural heritage protection and geological disaster risk mitigation.

2. Overview of the study area

2.1. Geological overview

The study area is located in the western part of Datong city, Shanxi province, on the western edge of the Datong–Lingqiu block of the North China Platform, as shown in Figure 1. The regional structure forms part of the Shanxi Graben Basin, with Cenozoic tectonic movements as the main controlling factor.[18] The stratigraphic sequence is relatively well-developed, consisting primarily of the Archean metamorphic rock series, Triassic red beds, Jurassic coal-bearing strata, and Quaternary loose deposits from bottom to top.[19] Among them, the Jurassic Datong Formation has a thick stratum and stable rock composition, making it the main coal-bearing stratum in this area, as demonstrated in Figure 2.

The coal-bearing strata in the study area are developed from bottom to top in multiple coal-bearing layers, with interlayers of mudstone, siltstone, and fine sandstone between the coal layers, exhibiting good stratified characteristics. In terms of structure, the overall formation is monoclinal, with the strata gently dipping toward the west at an angle of approximately 8–12°. No large-scale fold structures have been observed in the region.[20] Fractures are not well developed overall; however, due to the long-term impact of mining activities, joints and fractures in the shallow rock layers are relatively well developed. Localized areas of roof collapse and mining-induced fracture zones may pose potential threats to the overall stability of the rock mass.[21]

2.2. Coal seam distribution and current mining status

The exploitable coal seams in the study area mainly include coal seams #2, #3, #7, #8, #9, #11, and #12, all of which

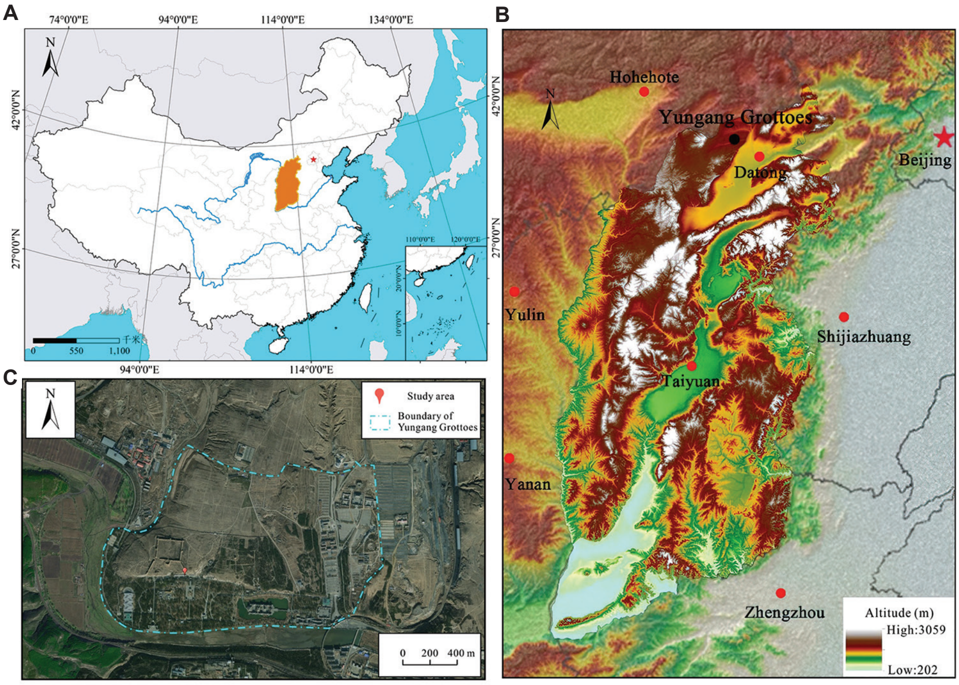


Figure 1. Research area map. (A) Shanxi province is located in North China. (B) The Yungang Grottoes are located in the northern part of Shanxi province at an average elevation of 1,186 m. (C) Yungang Grottoes study area.

occur within the Jurassic Datong coal-bearing strata.[22] There are variations in both thickness and depth among the coal seams. Seams #7, #9, and #11 are relatively stable and thick, serving as the main mining seams, as shown in Table 1. The thickness of coal seam #7 is approximately 2.36 m, buried at a depth of about 115 m. It is continuously deposited and of excellent quality, and large-scale mining has been completed. The thickness of coal seam #9 is 0.75 m, buried at a depth of approximately 125 m, and it has been the principal mining seam both currently and historically. The thickness of coal seam #11 is 1.57 m, buried at a depth of about 130 m. It is locally developed, and mining in some areas has been suspended. Other coal seams vary in thickness between 0.3 and 2.65 m, most of which are locally mineable. Due to long-term underground mining activities, numerous goafs have formed in the area. Some goafs exhibit collapse zones and fissure-type water-conducting zones at the top, while others contain significant water accumulation, posing considerable geological safety risks.

2.3. Geophysical characteristics

The surrounding rock medium in the study area mainly consists of mudstone, siltstone, and sandstone. Among these, sandstone typically has a higher wave velocity, with dense sandstone reaching wave velocities of over 1,200 m/s. Mudstone has a lower wave velocity, typically ranging from 750 to 950 m/s. Due to rock fragmentation, gas accumulation, or water accumulation, the wave velocity in goafs is significantly lower than that of the surrounding rock.[23,24] The specific wave velocity depends on the degree of fragmentation and the extent of water accumulation in the goaf.[25]

Electrical parameters also show significant differences. Normal surrounding rock has high resistivity, with siltstone and sandstone resistivities mostly ranging from 100 to 300 Ω·m, while mudstone resistivity is relatively low, approximately 50 to 150 Ω·m. In areas of water accumulation within goafs, the resistivity decreases markedly, usually to below 50 Ω·m, and locally to below 20 Ω·m. Such low-resistivity anomaly zones show strong spatial correlation with known water-filled goafs. The physical parameters of the Jurassic coal strata are shown in Table 2.

3. Research methods

3.1. Principle of passive-source surface wave exploration

SPAC is a passive surface wave detection technique based on microtremor observations. It is mainly used to extract Rayleigh wave dispersion curves and to invert underground shear wave velocity structures.[26]

The SPAC method is based on two assumptions: (i) micro-movements conform to a stationary

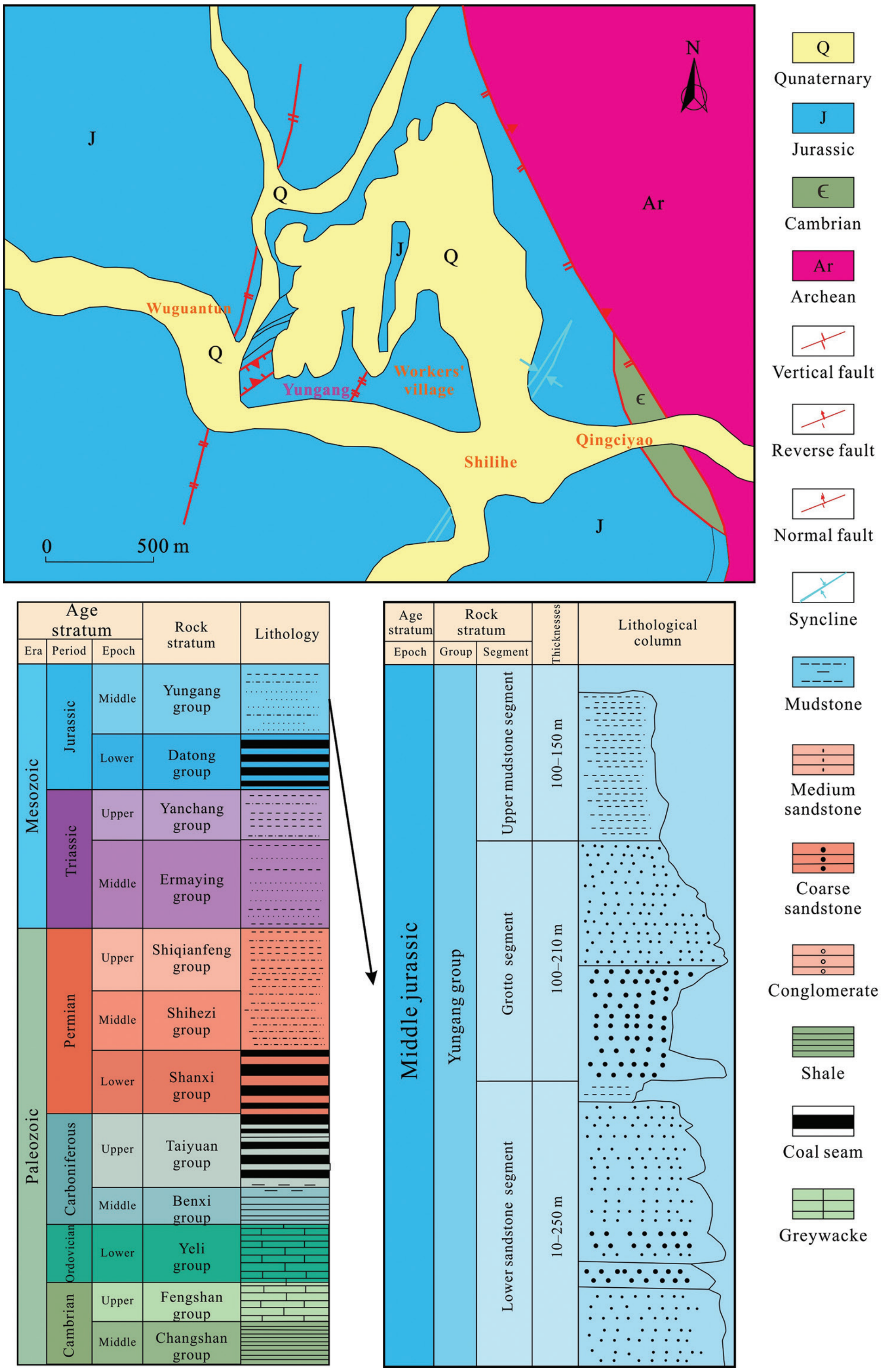


Figure 2. Geological map of Yungang Grottoes

random process in space and time, and (ii) among the various components contained in micro-movements, fundamental surface waves dominate. Let there be a reference observation point at the center point O(0,0), with a circle of observation points $A_1(r,\theta)$ distributed at a distance r around it, as shown in Figure 3. Then, the SPAC coefficient $\rho(\omega,r)$ between these two points can be expressed as **Equation (I)**:

Table 1. Characteristics of coal seams in the study area

Coal seam number	Coal seam minimum-maximum (average)/(m)	Top plate rock type	Base plate rock type
2^{-3}	1.45–6.12 (4.59)	Coarse sandstone; fine sandstone	Siltstone
3	0.10–8.82 (1.74)	Fine sandstone; siltstone	Siltstone
7^{-1}	0.00–5.50 (0.86)	Fine sandstone;	Carbonaceous mudstone
7^{-3}	0.27–4.79 (1.50)	Siltstone; fine sandstone	Sandy mudstone; Siltstone
8	0.00–3.00 (0.64)	Siltstone; fine sandstone	Siltstone; fine-grained sandstone
9	0.00–2.35 (0.75)	Fine sandstone; Sandstone	Sandy mudstone; Siltstone
11^{-1}	0.00–9.22 (1.27)	Fine sandstone; sandy mudstone	Siltstone; fine sandstone
11^{-3}	0.00–2.45 (0.30)	Sandy mudstone; siltstone	Sandy mudstone; fine sandstone
12^{-2}	0.00–9.22 (2.65)	Coarse sandstone; medium sandstone	Carbonaceous mudstone; sandy mudstone; siltstone
14^{-2}	0.00–4.67 (0.57)	Fine sandstone	Siltstone; fine sandstone; sandy mudstone

Table 2. Physical parameters of Jurassic coal strata

Rock type	Apparent density (g/cm^3)	Apparent resistivity (Ω·m)	Sound time difference (μs/m)	Wave velocity (m/s)
Coal seam	1.4–1.7 (1.6)	280–930 (640)	440–640 (542)	1,563–2,273 (1,845)
Mudstone	2.1–2.4 (2.3)	54–105 (80)	291–414 (351)	2,415–3,436 (2,849)
Sandy mudstone	2.2–2.6 (2.4)	69–121 (95)	280–400 (339)	2,500–3,571 (2,950)
Siltstone	2.3–2.7 (2.5)	72–130 (103)	241–350 (301)	2,857–4,149 (3,322)
Fine sandstone	2.4–2.7 (2.6)	85–151 (115)	211–312 (262)	3,205–4,739 (3,817)
Medium sandstone	2.5–2.8 (2.7)	108–390 (250)	172–243 (212)	4,115–5,814 (4,717)
Coarse sandstone	2.6–2.8 (2.7)	138–442 (297)	124–186 (161)	5,376–8,065 (6,211)
Gravel; rock	2.7–2.8 (2.7)	145–470 (308)	114–171 (145)	5,848–8,772 (6,897)

Note: Values in parentheses represent the average.

$$\rho(\omega,r)=\frac{1}{2\pi}\int_0^{2\pi}\varphi_\omega(r,\theta)\mathrm{d}\theta \qquad \text{(I)}$$

Among these, $\varphi_\omega(r,\theta)$ represents the coherence function at frequency ω, which can be expressed as **Equation (II)**:

$$\varphi_\omega(r,\theta)=\frac{u(O(0,0),t)\cdot u(A_1(r,\theta),t)}{|u(O(0,0),t)|\cdot|u(A_1(r,\theta),t)|} \qquad \text{(II)}$$

Here, $u(O(0,0),t)$ and $u(A_1(r,\theta),t)$ denote the microtremor signals at the center point and surrounding observation points at time t, respectively. Under isotropic conditions, the coherence of surface waves can be expressed using the zero-order Bessel function J_0 (**Equation [III]**):

$$\rho(\omega,r)=J_0\left(\frac{\omega r}{c(\omega)}\right) \qquad \text{(III)}$$

In this equation, J_0 is the zero-order Bessel function of the first kind, ω is the angular frequency, r is the distance between observation points, and $c(\omega)$ is the phase velocity at frequency ω.

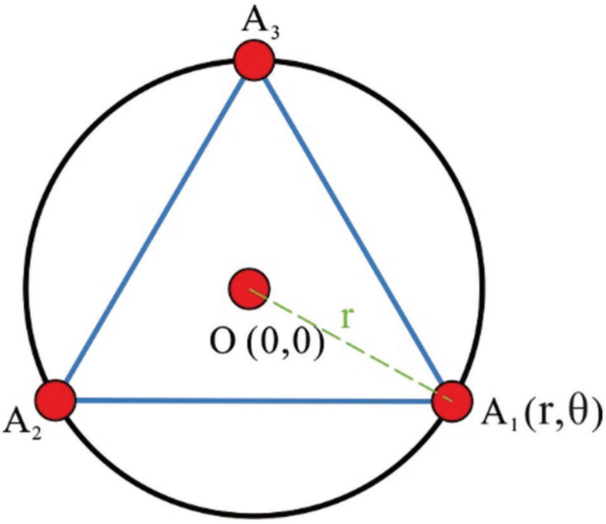


Figure 3. Single-circle observation array for the spatial autocorrelation method

By measuring the SPAC coefficient $\rho(\omega,r)$ at different distances r, the relationship between $\rho(\omega,r)$ and the Bessel function can be fitted, thereby deriving the phase velocity

$c(\omega)$ at different frequencies.[27] The curves showing the variation of these phase velocities with frequency are the surface-wave dispersion curves.[28]

The SPAC method is characterized by flexible deployment, high non-destructiveness, strong anti-interference ability, and adaptability to complex geological environments. It is particularly suitable for cultural heritage protection areas, as it does not rely on strong seismic sources and causes no disturbance to the ground surface. Therefore, it was employed in this study to detect the underground goaf in the Yungang Grottoes.

3.2. Shear wave velocity inversion

To invert the underground shear wave velocity structure from the dispersion curve, this study used a non-linear least-squares inversion algorithm and introduced a damping factor as a regularization term in the objective function to improve the stability of the inversion process and the physical plausibility of the model results.[29-31] The inversion objective function is expressed as follows (**Equation [IV]**):

$$\min_{m}\Phi(m)=\left\|d_{obs}-d_{cal}(m)\right\|^{2}+\lambda^{2}\left\|L(m-m_{0})\right\|^{2} \tag{IV}$$

Among these, m is the stratigraphic model parameter to be inverted (mainly shear wave velocity), d_{obs} and $d_{cal}(m)$ represent the measured and model-predicted dispersion data, respectively; λ is the damping factor; L is the regularization matrix; and m_0 is the prior model.

The damping factor λ in **Equation (V)** controls the trade-off between minimizing the data misfit and maintaining model smoothness. To systematically determine the optimal λ, the L-curve criterion[32] was employed, which constructs a curve by plotting $\log(\|d_{obs}\text{-}d_{cal}(m)\|)$ versus $\log(\|L(m\text{-}m_0)\|)$ for a series of candidate damping values. The optimal λ_{opt} corresponds to the corner of this L-shaped curve, where the curvature $k(\lambda)$ reaches its maximum (**Equation [V]**):

$$k(\lambda)=\frac{\left|\rho'(\lambda)\eta''(\lambda)-\rho''(\lambda)\eta'(\lambda)\right|}{\left[\rho'(\lambda)^{2}+\eta'(\lambda)^{2}\right]^{\frac{3}{2}}} \tag{V}$$

Figure 4. Ant-1C single-component nodal seismometer

Where $\rho(\lambda)=\log(\|d_{obs}\text{-}d_{cal}(m)\|)$, $\eta(\lambda)=\log(\|L(m\text{-}m_0)\|)$, and primes denote derivatives with respect to λ.

In implementation, 50 logarithmically spaced damping factors ranging from 10^{-3} to 10^{2} were tested based on preliminary sensitivity analyses. For each candidate λ, non-linear inversion was performed using the Levenberg–Marquardt algorithm with convergence criteria of relative change $<10^{-4}$ or a maximum of 100 iterations. The curvature $k(\lambda)$ was computed numerically from the resulting L-curve,[33] and the value maximizing $k(\lambda)$ was identified as the optimal balance point. The selected λ was validated by checking the geological plausibility of the inverted velocity model.

Introducing a damping factor effectively balances data-fitting accuracy and model complexity, avoiding unreasonable fluctuations in the model, especially under complex site conditions with high data noise or lateral heterogeneity in the subsurface structure.[34] Its regularization effect is crucial for improving the reliability of inversion results.[35] This study utilized parameter sensitivity testing to determine reasonable damping factor values and verified their positive effect in goaf identification, significantly improving the identification resolution of low-velocity anomalies such as cavities and fracture zones.

3.3. Equipment parameters

The node-type seismometer used in this experiment was the ANT-1C from Earth Pulse Technology Co., Ltd.

Table 3. ANT-1C single-component (5 Hz) nodal seismometer specifications

Parameter	Indicator (%)
Natural frequency	5±7.5
DC resistance	1850±5
Sensitivity	80±5
Damping	0.6±7.5
Distortion	≤0.1
Aliasing	≥170

Abbreviation: DC: Direct current.

(China) (Figure 4). This instrument is highly sensitive, has a wide dynamic range, and is well-suited for field use, making it ideal for long-term field observations. The device supports sampling intervals of 0.25/0.5/1.0/2.0/4.0 ms (1.0 ms, i.e., 1,000 sps, was used in this experiment), with a typical main frequency band of 1–100 Hz (the low frequency can be set to 1 Hz, and the upper high-frequency limit is limited by the sampling rate and anti-aliasing filter). The instantaneous dynamic range can reach 126.8 dB (@0 dB, 1 ms), and the total dynamic range can reach 150 dB. The equivalent input noise is 0.22 µV (12 dB gain, 2 ms). Internal storage is configurable from 8 GB to a maximum of 32 GB (this unit has 32 GB); time accuracy is ±10 µs, and the device includes a built-in global positioning system (BeiDou Navigation Satellite System, China National Space Administration, China) for timing and positioning. Detailed device parameters are provided in Table 3.

3.4. Technical optimization

To improve the quality of surface wave dispersion extraction and the resolution of shear wave velocity profiles, the layout and acquisition parameters were specifically optimized for this exploration.[36] A linear array configuration was adopted instead of the traditional equilateral triangular array structure. Linear arrays have been demonstrated to provide more efficient spatial sampling through $N(N\text{-}1)/2$ independent station pairs and better azimuthal coverage for directional ambient noise sources typical of urban heritage environments.[37] This geometry is particularly advantageous in spatially constrained sites where elongated deployment corridors (e.g., roadsides and pathways) are

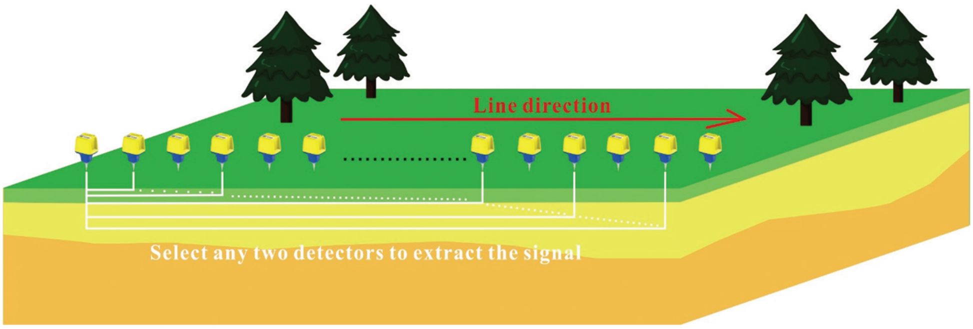


Figure 5. Extraction pattern diagram

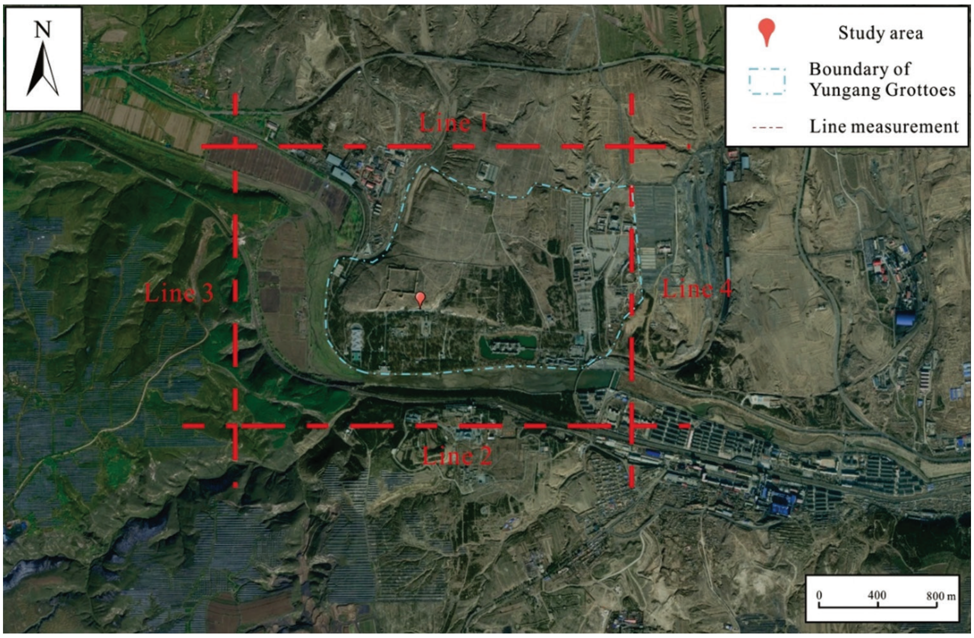


Figure 6. Schematic diagram of survey line arrangement

Table 4. Comparison with recent studies

Study	Array type	Damping strategy	Max depth (m)	Lateral res. (m)	RMS reduction	Application
This work	Linear, all-pairs	L-curve (0.20–0.35)	320	10–20	27%	Goafs, heritage
Ling *et al.*[9]	Circular	Fixed (0.15)	200	30–40	Not reported	Coal goafs
Li *et al.*[13]	2D grid	Tikhonov (0.20)	150	25–35	Not quantified	Fault imaging
Wang *et al.*[38]	Linear	None	180	40–50	N/A	Mine tunnels

Abbreviation: N/A: Not available; res.: Resolution; RMS: Root mean square.

Table 5. Statistics on surface wave exploration workload

Line arrangement	Length (m)	Coordinate points	Physical points
Line 1 (N)	3210	322	322
Line 2 (S)	3210	322	322
Line 3 (W)	2580	259	259
Line 4 (E)	2580	259	259

Abbreviations: E: East; N: North; S: South; W: West.

the only accessible options, as encountered in the Yungang Grottoes protection area.[38,39]

By adopting a processing method that extracts surface wave signals between any two stations (Figure 5), data information was fully utilized to achieve the highest possible spatial resolution in the imaging results. This combination of optimized array geometry and advanced signal processing effectively addresses the dual challenges of high-resolution subsurface imaging and non-invasive detection in sensitive cultural heritage environments.

3.5. Methodological innovations and performance metrics

This study introduced three key innovations that enhance imaging accuracy compared with conventional SPAC applications:

(i) Linear array with multi-pair processing: High-density linear arrays exploit $N(N\text{-}1)/2$ station pairs, yielding over 31,000 correlation pairs per survey line versus ~250 in conventional radial geometries.

(ii) L-curve-guided damping selection: Systematic curvature analysis across 50 candidates (10^{-3}–10^{2}) objectively determines λ_{opt}, eliminating subjective parameter selection.

(iii) Velocity-gradient-adaptive regularization: Optimized damping suppresses non-geological oscillations at sharp velocity contrasts ($\Delta v_s > 800$ m/s).

The following quantitative improvements were observed compared with conventional fixed-λ inversion: (i) root mean square residual reduction: average 27% (8 Hz: 42.3→29.6 m/s); (ii) oscillation suppression: 58% reduction in velocity coefficient of variation (0.34→0.14); (iii) anomaly contrast enhancement: 1.6:1→2.1:1; and (iv) resolution gain: 15 m @ 100 m depth versus 25 m.

Table 4 demonstrates the improved performance of the approach in this study compared with recent studies in maximum detection depth (320 m), lateral resolution (10–20 m), and RMS reduction (27%).

4. Experimental design and data processing

4.1. Measurement network layout

To obtain information on the shallow stratum shear wave velocity structure of the study area, four passive source surface wave lines, totaling 11.58 km in length, were laid out in the region, taking into account surface conditions and the distribution of underground safety red lines. A total of 1126 physical observation points were established, as depicted in Figure 6. The layout of the survey lines fully accounted for the distribution of goafs, surface accessibility, and acoustic signal coverage requirements. The overall layout was distributed in a north–south direction, following known underground safety control lines. The data are summarized in Table 5.

The spacing between each measurement point was moderate, with an average platform spacing of <10 m to ensure Rayleigh wave wavelength resolution while maintaining imaging accuracy. The on-site deployment employed a node-based cable-free connection, facilitating high-density deployment in densely built-up urban areas, along roadsides, or within cultural heritage protection zones, thereby effectively improving data integrity and spatial resolution.

4.2. Quality control

To ensure measurement accuracy and data quality, this experiment utilized a high-precision real-time kinematic-global navigation satellite system for real-time three-dimensional (3D) positioning of all node layout points. The horizontal positioning accuracy of the nodes was better than ±2 cm, and vertical elevation control accuracy was better than ±5 cm, meeting the strict requirements for station positioning in high-precision surface wave data inversion.

During data collection, a pass-rate verification mechanism and a real-time on-site monitoring system conducted multiple rounds of quality screening on the observed data. The physical point data in this study achieved a 100% pass rate. A total of 57 checkpoints were set up,

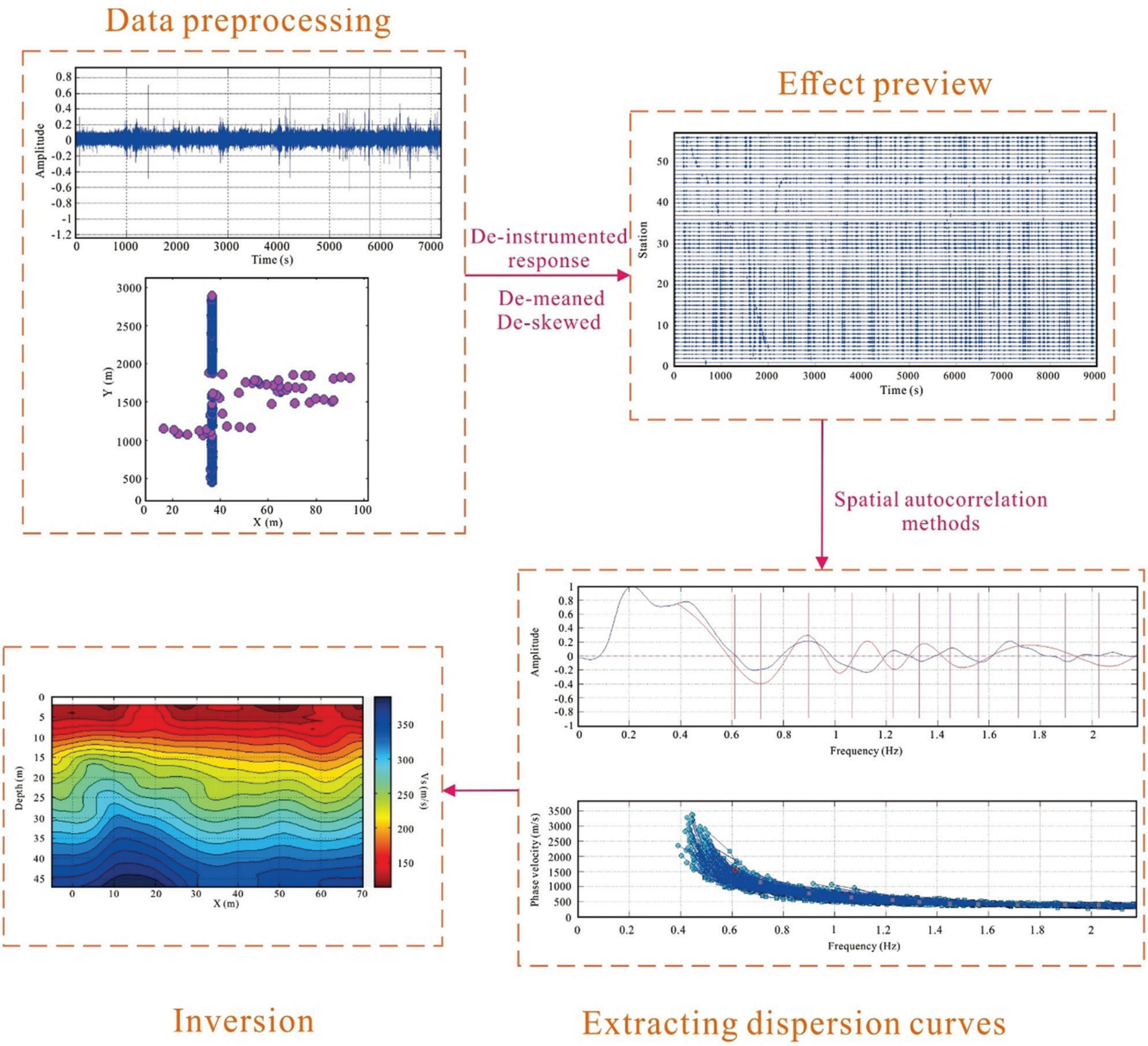


Figure 7. Flow chart of data processing

accounting for 5.02% of all measurement points, and each successfully passed full-waveform playback and spectrum analysis tests. No missed samples, incorrect recordings, or significant noise interference were detected, ensuring the authenticity and integrity of the data. Quantitative quality metrics show signal-to-noise ratio ranging from 15 to 45 dB across frequency bands of 2–15 Hz, with an average waveform correlation coefficient >0.85 between adjacent stations, confirming data reliability for dispersion analysis.

4.3. Data processing flow

After data acquisition, the microtremor data were processed and inverted using tomography. The specific process includes four stages: data preprocessing, dispersion analysis, velocity inversion, and anomaly interpretation. During processing, the original microtremor waveform was first subjected to noise filtering and waveform shaping, followed by the extraction of the dominant Rayleigh wave component to construct a frequency-phase velocity spectrum.[40] The dispersion curve extraction used multi-channel spectrum superposition technology to effectively suppress noise interference. The transverse wave velocity profile inversion utilized a multi-frequency co-inversion algorithm, supplemented by tomography technology to construct a 2D wave velocity structure, as shown in Figure 7.The specific method and process are as follows: using continuous waveform data recorded by the array, the

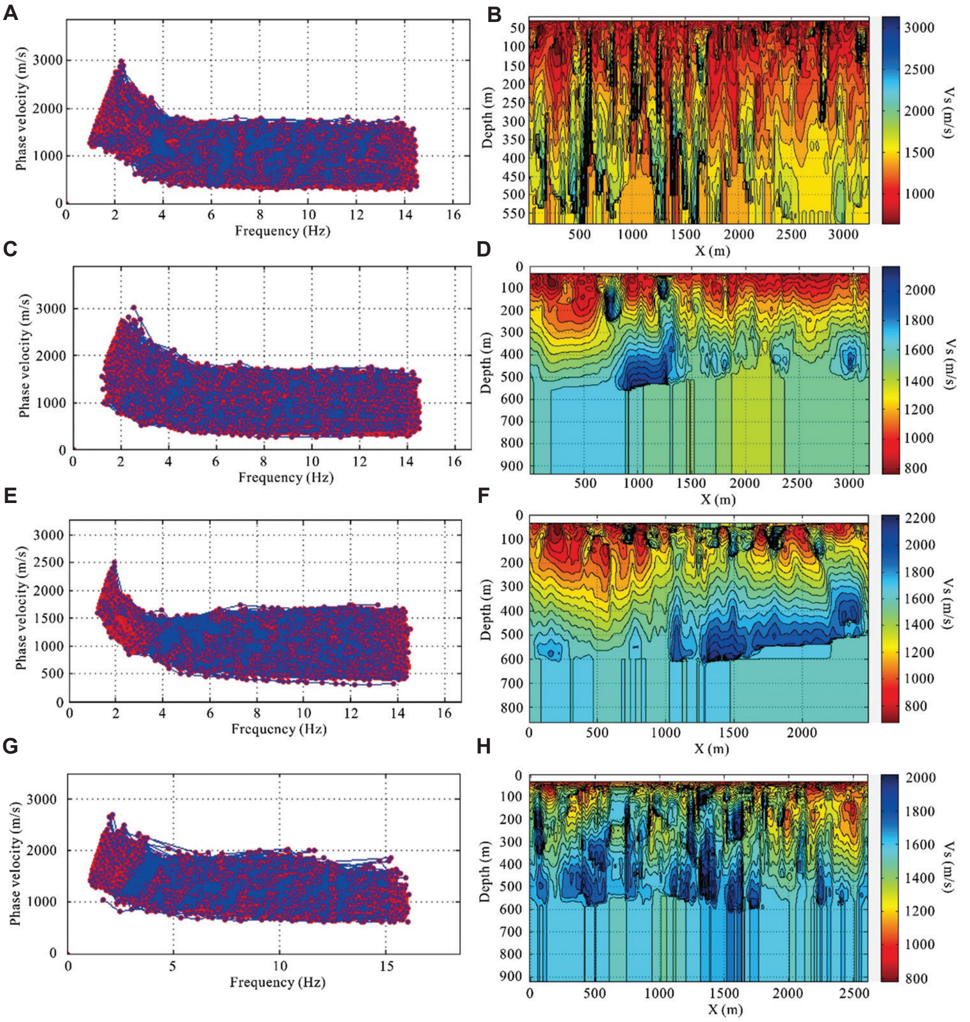


Figure 8. Dispersion curves and shear wave velocity profiles. (A) and (B) show the dispersion curves and two-dimensional velocity profiles for Line 1; (C) and (D) for Line 2; (E) and (F) for Line 3; and (G) and (H) for Line 4.

frequency domain waveform SPAC method was applied to obtain the surface wave Green's function between stations. The processing techniques include raw data preprocessing, such as instrument response removal, mean removal, tilt removal, and segmentation into unit-length data (e.g., minutes). Seismic signal interference was suppressed through preprocessing methods like sliding absolute value averaging. The processed waveforms were then spectrally whitened within the required frequency band. Next, the SPAC function between two stations was calculated every 2 min. Finally, the long-term autocorrelation function was superimposed as required, allowing further extraction of the surface wave dispersion curve.[41] For extracting surface wave dispersion curves, the SPAC method was used to derive surface wave phase velocity dispersion curves, while surface wave tomography obtained dispersion curves for each point beneath the array. Specifically, the inversion region was first divided into a grid, and the phase velocity distribution for each cycle is inverted to ultimately obtain a surface wave apparent velocity structure model of the medium beneath the array.[42]

4.4. Data imaging results

The survey consisted of four measurement lines, each comprising multiple seismic arrays and nodes strategically deployed to capture detailed subsurface shear wave velocity information. The specifics of each line's layout and data acquisition are as follows:

(i) Line 1 consists of two observation arrays. Array 1 was completed by 239 seismic stations at points 1083–1322, while Array 2 was completed using 246 seismic stations at points 1001–1246.

(ii) Line 2 also comprises two observation arrays. Array 1 includes 246 seismic stations at points 2001–2246,

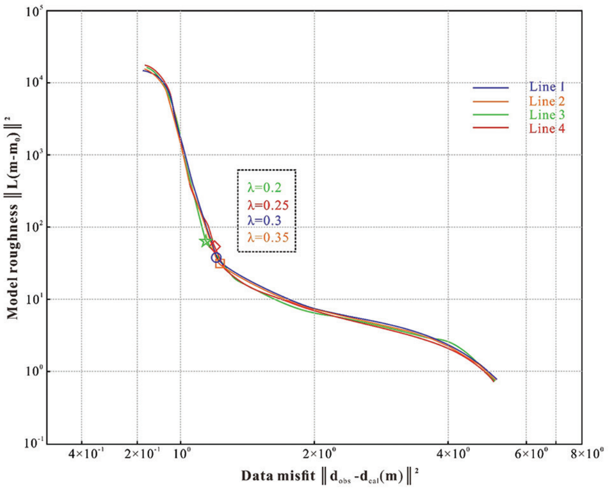


Figure 9. L-curve analysis for determining the optimal damping factors of the four survey lines. Notes: This figure presents the L-curve analysis used to select the optimal damping factors for the four survey lines. The horizontal axis represents the model norm ($\|d_{obs}\text{-}d_{cal}(m)\|$), while the vertical axis represents the residual norm ($\|L(m\text{-}m_0)\|$). Each curve shows the trade-off between model complexity and data fit at different damping factor values. The inflection point of each L-curve corresponds to the optimal damping factor, which balances model complexity and data misfit. The selected damping factors, located near the inflection points, ensure the stability and physical reasonableness of the inversion results.

and Array 2 includes 236 seismic stations at points 2076–2312.

(iii) Line 3 was completed with 246 nodes positioned between observation points 3008 and 3254, along with one seismograph.

(iv) Line 4 involved a total of 244 nodes from measurement points 4001 to 4259, observed by one seismograph, completing one observation.

The complete imaging results of these four lines are illustrated in Figure 8.

To determine the optimal damping factor for the regularized inversion, this study employed the L-curve criterion to quantitatively analyze the inversion results of four survey lines (Figure 9). The L-curve is a plot in double logarithmic coordinates, with the model norm ($\|d_{obs}\text{-}d_{cal}(m)\|$) on the horizontal axis and the residual norm ($\|L(m\text{-}m_0)\|$) on the vertical axis, generated by inverting a series of damping factors λ.[43] This curve reflects the trade-off between model complexity and data fitting accuracy: smaller λ values yield better data fitting but can produce overly complex models that amplify noise, whereas larger λ values enhance model stability at the cost of reduced spatial resolution and excessive smoothing. The inflection point of the curve is considered the parameter that provides the best compromise between noise suppression and model resolution, that is, the optimal balance between the two.[44]

As shown in Figure 9, the L-curves of all four measurement lines exhibit the characteristic "L" shape with clearly identifiable inflection points. The damping factors selected based on this criterion are all near the inflection points, indicating a successful balance between stability of the results and physical plausibility during inversion. This approach reduces subjective bias in parameter selection and provides quantitative evidence for the reliability of subsequent inversion results.

The final damping factors ranged between 0.20 and 0.35. By adjusting the damping parameters within this range, the inversion model effectively balanced fitting accuracy with stability, thereby improving the reliability of identifying low-velocity anomalous zones. Figure 10 compares the imaging results before and after damping factor adjustment, showing that detection depth covers a range from 50 to 320 m, with lateral resolutions of approximately 10–20 m. The effects of seismic station distribution and terrain undulations are also well captured in the imaging.

4.5. Resolution verification

Synthetic checkerboard tests ($\Delta v_s = \pm 20\%$, 15 m × 20 m spacing) validated the stated resolution using Table 2 velocities and λ=0.25. The results were as follows: (i) 15 m features were resolved at depth <150 m (correlation >0.75); (ii) 20 m features were resolved at 150–250 m (correlation >0.70); and (iii) detectability extended to 320 m for ±15% contrasts.

A synthetic goaf (40 m wide, 8 m thick, v_s=1,000 m/s at 120 m) was recovered with <5 m horizontal and <3 m vertical errors, validating the observed goaf dimensions (50–200 m extent, 2–10 m thickness).

5. Results analysis and validation

5.1. Goaf identification mark

Low-velocity anomalies were identified through shear wave velocity inversion profiles, providing an effective physical parameter basis for goaf identification.[45,46] Research demonstrates that goafs exhibit significantly low-velocity characteristics in wave velocity imaging, with shear wave velocities generally below 1,200 m/s, clear boundary contours, and mostly band-shaped or lens-shaped anomalies. This velocity threshold, combined with on-site lithology, engineering history, and mine area data analysis, is used as the main geophysical indicator for determining the existence of goafs.[47-49]

Multi-level goafs often manifest as sudden drops in vertical velocity accompanied by increased velocity fluctuations, while horizontally they appear as continuous

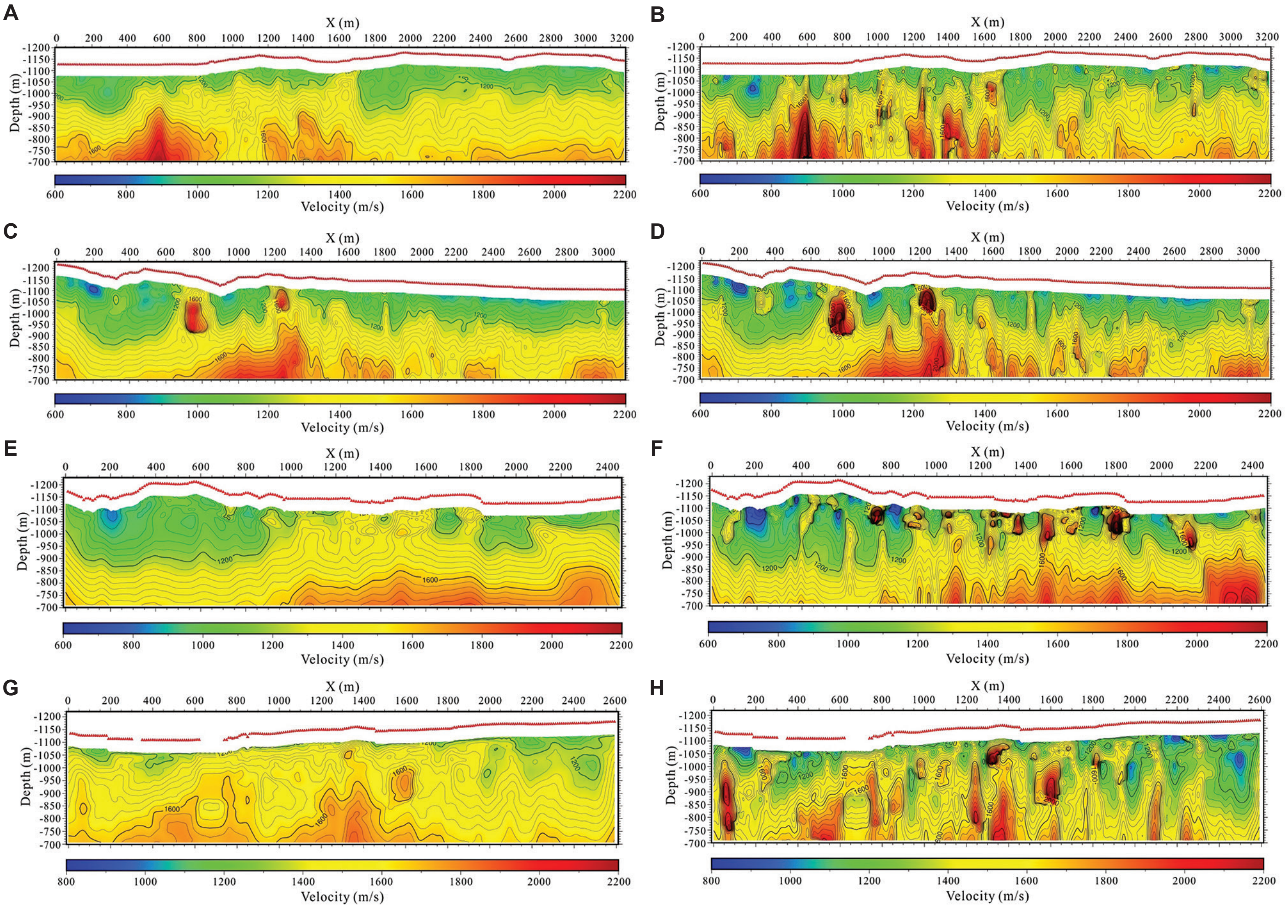


Figure 10. Comparison before and after adjustment of the damping factor. (A) Line 1, unadjusted damping factor. (B) Line 1, adjusted damping factor. (C) Line 2, unadjusted damping factor. (D) Line 2, adjusted damping factor. (E) Line 3, unadjusted damping factor. (F) Line 3, adjusted damping factor. (G) Line 4, unadjusted damping factor. (H) Line 4, adjusted damping factor.

low-velocity zones with wave velocity gradients significantly weaker than those of the surrounding normal rock. These characteristics provide parameter support for subsequent anomaly boundary delineation and stratification analysis.

5.2. Distribution pattern analysis

In this study, through systematic analysis of the wave impedance tomography results from four survey lines (L1 – N line, L2 – S line, L3 – W line, and L4 – E line), multiple typical low-velocity anomaly zones were identified. Combined with historical mining data, borehole records, and topographical information, their causes were accurately determined, as shown in Figure 11. The distribution of these anomalies is highly correlated with historical mining activities in the region, mainly manifested in a significant decrease in the apparent wave velocity at the depth of the coal seam and a concave wave velocity contour line, which are typical geophysical response characteristics of goafs.

Line 1 runs along the north side of the Yungang Grottoes and crosses several historical mining areas. The imaging results showed multiple coal seams with overlapping goafs along this line. Among them, the area between pile numbers 400–440 m (Line 1-1) is located beneath the riverbed of the Shilihe River, where wave velocity anomalies were significant. Based on historical data, this anomaly is inferred to be caused by the mining of coal seams #2, #3, and #7. The section between pile numbers 610 and 1,530 m (Line 1-2 to Line 1-4) also demonstrated pronounced low-velocity anomalies, which coincide with the mining activities of coal seams #2 and #3 at the Wuguantun Mine. The section between pile numbers 1,700 and 2,000 m (Line 1-5) is an area where multiple coal seams were jointly mined in the history of the Wuguantun Mine. The mining seams include #2, #7, #11-3, and #12. In the 2,140–2,930 m section (Line 1-6), low velocity anomalies extend deeply into coal seams #2, #3, and #7, reflecting multi-level mining characteristics.

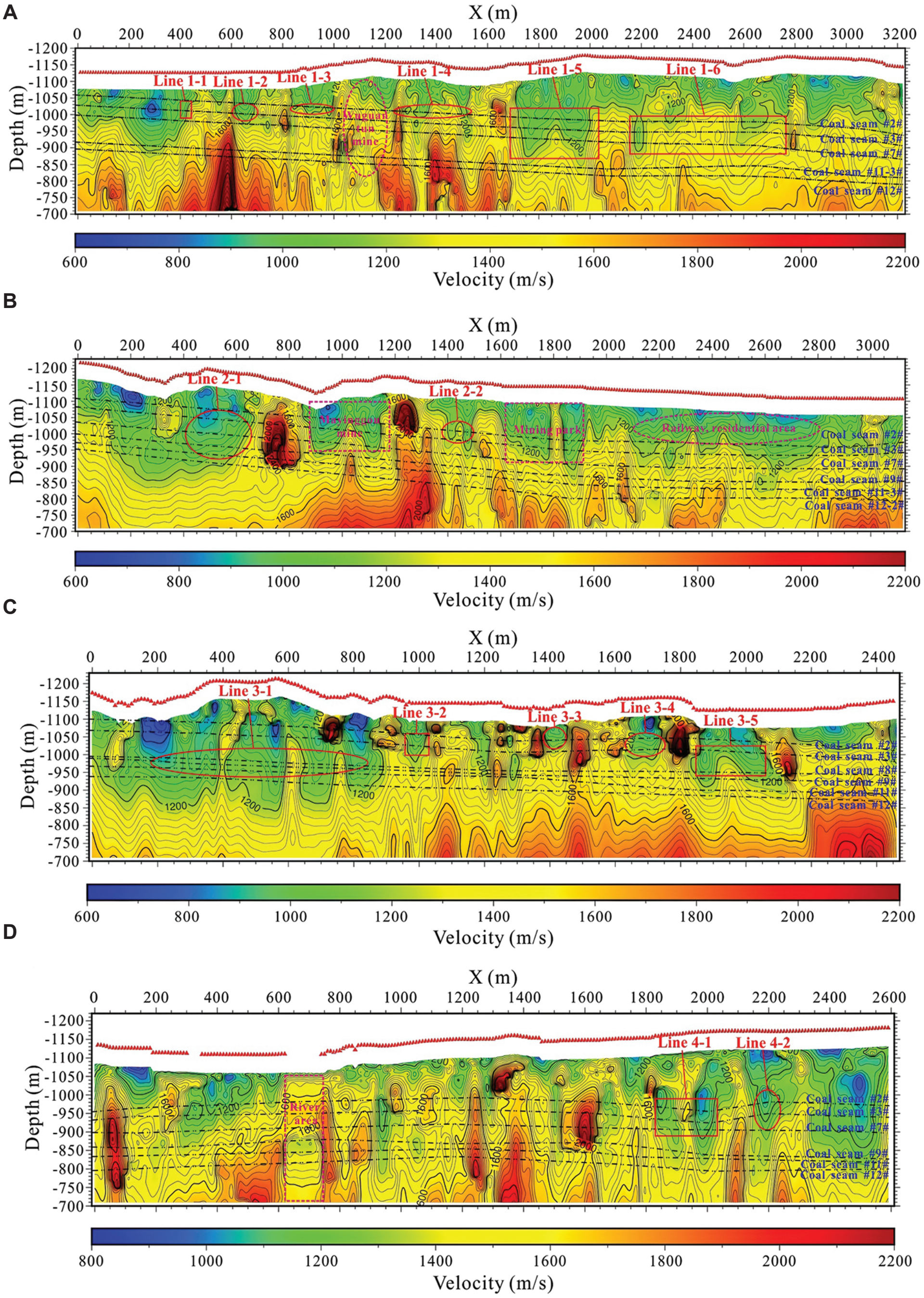


Figure 11. Transverse wave velocity profile interpretation diagram. (A) Line 1 two-dimensional (2D) velocity profile interpretation. (B) Line 2 2D velocity profile interpretation. (C) Line 3 2D velocity profile interpretation. (D) Line 4 2D velocity profile interpretation.

Line 2 is located on the south side of the Yungang Grottoes. The layer imaging results showed that multiple coal seams simultaneously exhibit significant reductions in wave velocity in the vertical direction, which is a typical feature of multi-layer mining. The section between pile numbers 400 and 660 m (Line 2-1) is a historically known goaf area of coal seams #9 and #12. The imaging results also indicate that the lower coal seams #2 and #3 experience a decrease in wave velocity and contour line disturbances. Combined with drilling data and mining records, this is inferred to be a multi-level superimposed goaf area. In addition, an isolated low-velocity anomaly occurs in the section between pile numbers 1,370 and 1,500 m (Line 2-2), for which no clear mining records exist. It is speculated that this anomaly results from a hidden void structure caused by small-scale mining.

Line 3 runs along the west side of the Yungang Grottoes. Within the range of pile numbers 320–920 m (Line 3-1), the results correspond to the mining records of coal seams #7–3, #9, and #12. The tomography depicts consistently low wave velocity and continuous structure, confirming the impact of multiple-seam mining. At 950–1,020 m (Line 3-2), a small kiln mining zone exhibits obvious wave velocity anomalies. In addition, three low-velocity anomaly zones (Line 3-3 to Line 3-5) are distributed in the Shilihe riverbed section (pile numbers 1380–2060 m), characterized by significant wave velocity concavity and the development of a fracture zone in the deeper parts of coal seams #2 and #3.

Line 4 is laid out on the east side of the Yungang Grottoes. The imaging results revealed two significant low-velocity anomaly zones distributed between pile numbers 1,830 and 2,250 m. In Line 4-1 (1,830–2,030 m), the regional wave impedance image shows low-velocity anomalies in the deeper coal seams, which are presumed to be goafs formed by the joint mining of coal seams #2 and #3 in small kilns. In the Line 4-2 section (2,160–2,250 m), the wave velocity ranges between 1,000 and 1,200 m/s, typical of the #7-1 coal seam mining reflection characteristics. At the same time, the #2 and #3 coal seams also display abnormal responses, indicating the existence of multiple layers of superimposed mining.

5.3. Summary of achievements

Through systematic and standardized field geophysical exploration, data processing, and comprehensive interpretation, combined with the results of field investigations, the distribution range of the main coal seam goafs below the geophysical survey lines in the protected area has been basically identified. When analyzing the extent of the goafs, mining data collected from the three coal mines surrounding the Yungang Grottoes were fully integrated. Finally, through comparative analysis and integration with known hydrogeological data, the extent of the goafs of each coal seam along the underground safety line of the Yungang Grottoes was delineated.

The extent of the goafs delineated along each survey line is described below:

(i) L1 survey line
  Line 1-1 is inferred to represent small kiln mining voids in coal seams #2 and #3, with a 300 m range located outside the boundary and a 40 m range located on the Yungang Grottoes underground safety line. Line 1-2 is presumed to represent small mine workings in coal seams #2 and #3, with an area of 80 m. Line 1-3 is also presumed to represent small mine workings in coal seams #2 and #3, with an area of 170 m. Line 1-4 is inferred to be a small kiln mining void in coal seams #2 and #3, with a range of 260 m. Line 1-5 is inferred to be a multi-layer goaf of coal seams #2, #3, #7, and #11, with a range of 340 m. Line 1-6 is presumed to be a multi-layer goaf of coal seams #2, #3, and #7-3, with a range of 870 m, of which 170 m lies outside the boundary.

(ii) L2 survey line
  Line 2-1 is inferred to represent multiple goafs of coal seams #2, #3, #7-1, #9, and #12, with a total goaf range of 670 m, of which 270 m lies on the underground safety line of the Yungang Grottoes, corresponding to the goaf of coal seam #9. The 400 m range located outside the boundary includes the 0–210 m section corresponding to the goaf of the #7-1 coal seam working face, and the 210–400 m section corresponding to the multi-layer goaf of the #7-1, #9, and #12 coal seams. Line 2-2 is presumed to represent small kiln mining voids in coal seams #2 and #3, with a range of 800 m, of which 620 m corresponds to the current surface of the mining park.

(iii) L3 survey line
  Line 3-1 represents a comprehensive inference of multiple goafs in coal seams #2, #3, #7-1, #9, and #12, with a total range of 980 m, of which 580 m lies on the underground safety line of the Yungang Grottoes. The 400–520 m section corresponds to the multi-layer goaf of coal seams #7-1, #9, and #12, the 520–910 m section corresponds to the multi-layer goaf of coal seams #7-1 and #9, and the 910–980 m section corresponds to the goaf of coal seam #7-1. Line 3-2 is inferred to be a multi-layer goaf of coal seams #2, #3, and #7-1, with a range of 80 m. Line 3-3 is presumed to be the goaf of coal seams #2 and #3, with a range of 80 m. Line 3-4 is presumed to be the goaf of small kilns in coal seams #2 and #3, with a range of 127 m. Line 3-5 is presumed to be the goaf of small kilns in coal seams #2 and #3, with a range of 220 m.

(iv) L4 survey line
  Line 4-1 is inferred to be a small kiln goaf in coal seams #2 and #3, with a range of 225 m. Line 4-2 is inferred to be a multi-layer goaf in coal seams #2, #3, and #7, with a range of 300 m, of which 70 m lies on the underground safety line of the Yungang Grottoes.

A total of 15 anomalous goafs were identified, including eight historical goafs from small coal mines. The anomalies are widely distributed along the L1 to L4 survey lines, showing a multi-center distribution pattern with densely developed goafs in local sections. This reflects the multi-level superimposed effects formed by coal mining activities during different historical periods. Lines 1 and 4 exhibit a high concentration of mining anomalies, characterized by large areas of low wave velocity, significant depth extension, and complex structural characteristics.

The goafs of small local coal mines are scattered and isolated, mostly concentrated near the intersection of survey lines and areas with uneven terrain, which may be related to early manual mining methods and terrain control. Although such small goafs are small in volume, they have a significant impact on geological structural integrity and engineering safety.

5.4. Implications for grotto stability

The identified low-velocity anomalies present varying geological risks depending on their spatial distribution, depth, and velocity characteristics. Four anomalies within the safety boundary (Line 1-1: 40 m, Line 2-1: 270 m, Line 3-1: 580 m, and Line 4-2: 70 m), located at depths of 50–180 m, pose potential subsidence risks.[50] Multi-layer goafs, such as L2-1, significantly increase the likelihood of progressive failure and accelerate weathering processes.

A shear wave velocity (v_s) below 1,200 m/s indicates highly fractured rock compared with intact rock (v_s>2,000 m/s). Using the empirical relationship rock quality designation (RQD)≈$0.01\times v_s^{1.35}$, a velocity of 1,200 m/s yields an RQD of approximately 28%, classifying the rock mass as "very poor" with low load-bearing capacity.[51]

The distribution of goafs also leads to stress concentrations 2–3 times higher than the lithostatic pressure,[52] increasing the risks of progressive roof collapse, surface fissuring in shallow goaf areas (<100 m depth), and asymmetric loading on structural pillars. Areas with water accumulation, such as Line 3-3 to Line 3-5, exhibit v_s<1,000 m/s, suggesting weakened rock strength, increased weathering, and susceptibility to freeze–thaw damage.

Risk zoning can be categorized as follows: (i) high-risk zones, within 50 m of the safety boundary, with v_s<1000 m/s and multi-layer goafs, require immediate monitoring and possible grouting; (ii) moderate-risk zones, between 50–150 m with v_s between 1,000–1,200 m/s, should undergo periodic surveys; and (iii) low-risk zones, beyond 200 m with isolated small goafs, should be monitored over the long-term.

Recommended monitoring techniques include global navigation satellite system surface deformation measurements, repeat SPAC surveys every 2–3 years, groundwater level monitoring, and InSAR for regional subsidence detection.[53]

5.5. Methodological transferability and limitations

The SPAC method is highly transferable to cultural heritage sites facing subsurface hazards, such as mining-affected monuments (e.g., Longmen and Mogao), karst sites,[54,55] and urban heritage zones with voids. Its non-invasive, ambient-noise-based operation aligns with conservation protocols, making it suitable for sensitive environments where traditional geophysical methods are restricted.

Key prerequisites include sufficient ambient noise, station spacing ≤0.3 × depth, and favorable geological conditions (avoiding high water tables >5 m). For shallow depths (<100 m), 5 Hz sensors are sufficient, while broadband sensors are required for deeper targets (>200 m). The optimal damping factor (λ=0.20–0.35), selected using L-curve analysis, is suitable for moderate velocity contrasts.

Limitations include electromagnetic interference (1–5 Hz), ambiguity in distinguishing natural and anthropogenic voids (both exhibiting v_s<1,200 m/s), and depth trade-offs (small voids <5 m at depths >150 m may require ground-penetrating radar or microgravity surveys).

Verification can be achieved through drilling (Line 1-1, Line 2-1, Line 3-1, and Line 4-2), ERT for water-filled voids (<50 Ω·m), microgravity surveys with 5 m spacing, repeat SPAC/InSAR (Δv_s<−5%), and 3D modeling for engineering-grade characterization.

6. Conclusion

Based on the urgent need for detailed detection of underground goafs in the Yungang Grottoes Protection Area, this study introduced the SPAC method combined with damping-optimized inversion to carry out systematic geophysical detection. The following main conclusions were reached:

(i) High-density linear arrays achieved 10–20 m lateral resolution and a maximum depth of 320 m, overcoming traditional imaging limitations in spatially constrained heritage sites.

(ii) Damping regularization (λ = 0.20–0.35) effectively suppresses velocity oscillations at sharp contrasts, improving model stability and goaf identification accuracy.

(iii) Fifteen low-velocity anomalies (v_s<1200 m/s) were identified, corresponding to historical mining of coal seams 2#, 3#, 7#, 9#, and 12#. Critical anomalies beneath the grotto core area pose significant geological hazard risks.

(iv) The non-invasive SPAC method achieves high-precision imaging without surface disturbance or artificial sources, demonstrating strong engineering adaptability for cultural heritage protection.

(v) The results provide a scientific basis for stability assessment and risk management, supporting heritage protection planning, goaf treatment, and safety supervision.

(vi) The framework is applicable to mining temples, karst sites, urban zones with sufficient noise, station spacing ≤0.3 × depth, and favorable geological conditions. L-curve damping (0.1–0.5) is universally applicable, while site-specific calibration optimizes performance.

(vii) Limitations include inversion non-uniqueness; ambiguity in interpreting v_s<1,200 m/s (needing borehole or ERT verification), near-surface bias (<30 m), and resolution degradation >200 m. The damping assumption of smoothness; abrupt contrasts need tuning. Multi-method integration is essential.

Overall, SPAC-damping optimization effectively achieves 10–20 m horizontal resolution and 320 m depth imaging without excavation and is transferable to World Heritage Sites for evidence-based risk management.

Acknowledgments

None.

Funding

This research was funded by the National Key Research and Development Project (grant number 2022YFC2903402).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Jiale Liu, Xiaodong Wang
Formal analysis: Jiale Liu, Xiaodong Wang
Investigation: Jiale Liu, Yanhai Liu, Ke Ren, Shangqing Zhang
Methodology: Jiale Liu, Xiaodong Wang
Visualization: Jiale Liu
Writing–original draft: Jiale Liu
Writing–review and editing: Jiale Liu, Xiaodong Wang, Yanhai Liu, Ke Ren

Availability of data

Due to confidentiality constraints, the data analyzed in this study are not available for public dissemination.

References

1. Vadrucci M. Sustainable cultural heritage conservation: A challenge and an opportunity for the future. *Sustainability*. 2025;17(2):584.

   doi: 10.3390/su17020584

2. Huang J, Ren J. Geophysical prospecting method applied in the preservation of Yungang Grottoes. *Sci Conserv Archaeol*. 2011;23(2):87-95.

   doi: 10.16334/j.cnki.cn31-1652/k.2011.02.007

3. Galdón JM, Rey J, Martínez J, Hidalgo MC. Application of geophysical prospecting techniques to evaluate geological-mining heritage: The Sinapismo mine (La Carolina, Southern Spain). *Eng Geol*. 2017;138:71-78.

   doi: 10.1016/j.enggeo.2017.01.012

4. Martinho E, Dionísio A. Main geophysical techniques used for non-destructive evaluation in cultural built heritage: A review. *J Geophys Eng*. 2014;11(5):053001.

   doi: 10.1088/1742-2132/11/5/053001

5. Alao JO, Lawal KM, Dewu BBM, Raimi J. Construction of a multi-purpose geophysical test site on a lateritic clay soil. *Arab J Geosci*. 2024;17:238.

   doi: 10.1007/s12517-024-12039-72

6. Alao JO, Lawal KM, Dewu BBM, Raimi J. Near-surface seismic refraction anomalies due to underground target models and their application in civil and environmental engineering. *Phys Chem Earth Parts A/B/C*. 2025;138:103845.

   doi: 10.1016/j.pce.2024.103845

7. Perski Z, Hanssen R, Wojcik A, Wojciechowski T. InSAR analyses of terrain deformation near the Wieliczka Salt Mine, Poland. *Eng Geol*. 2009;106(1-2):58-67.

   doi: 10.1016/j.enggeo.2009.02.014

8. Xu PF, Li CJ, Ling SQ, Zhang YB, Hou Z, Sun YJ. Mapping collapsed columns in coal mines utilizing Microtremor Survey Methods. *Chin J Geophys*. 2009;52(7):1923-1930.

   doi: 10.3969/j.issn.0001-5733.2009.07.028

9. Ling S, Ren Q, Cheng F, *et al*. Application of microtremor survey technology in a coal mine goaf. *Appl Sci*. 2023;13(1):466.

   doi: 10.3390/app13010466

10. Martínez-Moreno FJ, Pedrera A, Ruano P, *et al*. Combined microgravity, electrical resistivity tomography, and induced polarization to detect deeply buried caves: Algaidilla cave (Southern Spain). *Eng Geol*. 2013;162:67-78.

    doi: 10.1016/j.enggeo.2013.05.008

11. Metwaly M, AlFouzan F. Application of 2-D geoelectrical resistivity tomography for subsurface cavity detection in the eastern part of Saudi Arabia. *Geosci Front*. 2013;4(4):469-476.

    doi: 10.1016/j.gsf.2012.12.005

12. Styles P, McGrath R, Thomas E, Cassidy NJ. The use of microgravimetry for cavity characterization in karstic terrains. *Q J Eng Geol Hydrogeol*. 2005;38(2):155-169.

    doi: 10.1144/1470-9236/04-035

13. Li Q, Zhang H, Lei X, Li C. Imaging of upper breakpoints of buried active faults through microtremor survey technology. *Earth Planets Space*. 2024;76:132.

    doi: 10.1186/s40623-024-02080-x

14. Liu Y, Xia J, Guan B, Xi C, Ning L, Zhang H. Short-term synchronous and asynchronous ambient noise tomography in urban areas: Application to karst investigation. *Engineering,* 2025;48:292-308.

doi: 10.1016/j.eng.2025.02.001

15. Alao JO, Lawal KM, Dewu BBM, Raimi J. Detection of shallow underground targets using electrical resistivity tomography and the implications in civil/environmental engineering. *Discov Geosci*. 2024;2:52.

doi: 10.1007/s44288-024-00058-6

16. Zeid AN, Corradini E, Bignardi S, Nizzo V, Santarato G. The passive seismic technique 'HVSR' as a reconnaissance tool for mapping paleo-soils: The case of the Pilastri archaeological site, Northern Italy. *Archaeol Prospect*. 2017;24(3):245-258.

doi: 10.1002/arp.1568

17. Schwellenbach I, Hinzen K, Petersen MG, Bottari C. Combined use of refraction seismic, MASW, and ambient noise array measurements to determine the near-surface velocity structure in the Selinunte Archaeological Park, SW Sicily. *J Seismol*. 2020;24:753-776.

doi: 10.1007/s10950-020-09909-4

18. Huang X, Guo F, Zhao L, Zhang F. Chemical weathering records in Yungang Formation, North China: Implications for stone heritage conservation. *Front Earth Sci*. 2025;13:1507580.

doi: 10.3389/feart.2025.1507580

19. Zhang Y, Shi W, Dong S, Wang T, Yang Q. Jurassic intracontinental deformation of the central North China Plate: Insights from syn-tectonic sedimentation, structural geology, and U-Pb geochronology of the Yungang Basin, North China. *Tectonophysics*. 2020;778:228371.

doi: 10.1016/j.tecto.2020.228371

20. Liu A, Xu Y, Liu C, Pang E. Geological characteristics and tectonic evolution of Datong Basin. *Geoscience*. 2021;35(5):1296-1310.

doi: 10.19657/j.geoscience.1000-8527.2020.067

21. Chen B, Yu X, Wang T, Ma F, Li S, Yang L. Lithofacies and architectural characteristics of sandy braided river deposits: A case from outcrops of the Middle Jurassic Yungang Formation in the Datong Basin, Shanxi Province. *Oil Gas Geol*. 2015;36(1):111-117.

doi: 10.11743/ogg20150114

22. Huang J. Study on the geological characteristics of the Yungang Grottoes. *Southeast Cult*. 2003;5:91-93.

23. Guo Y, Qin Y, Chen P, Xu N. Simulation of the compaction behavior and the water permeability evolution of broken rock masses of different shapes in a goaf. *Water*. 2023;15(6):1190.

doi: 10.3390/w15061190

24. Berryman JG. Seismic waves in rocks with fluids and fractures. *Geophys J Int*. 2007;171(2):954-974.

doi: 10.1111/j.1365-246X.2007.03563.x

25. Yang H, Duan HF, Zhu J, Zhao Q. Water effects on elastic S-wave propagation and attenuation across single clay-rich rock fractures: Insights from ultrasonic measurements. *Rock Mech Rock Eng*. 2024;57:2645-2659.

doi: 10.1007/s00603-023-03712-6

26. Aki K. Space and time spectra of stationary stochastic waves with special reference to microtremors. *Bull Earthq Res Inst Univ Tokyo*. 1957;35:415-456.

27. Xu P, Li S, Ling S, Guo H, Tian B. Application of SPAC method to estimate the crustal S-wave velocity structure. *Chin J Geophys*. 2013;56(11):3846-3854.

doi: 10.6038/cjg20131126

28. Xu P, Ling S, Li C, *et al*. Mapping deeply-buried geothermal faults using microtremor array analysis. *Geophys J Int*. 2012;188(1):115-122.

doi: 10.1111/j.1365-246X.2011.05266.x

29. Tarantola A, Valette B. Inverse problems = Quest for information. *J Geophys*. 1981;50(1):159-170.

30. Tikhonov AN. Solution of incorrectly formulated problems and the regularization method. *Sov Math Dokl*. 1963;5:103-114.

31. Phillips DL. A technique for the numerical solution of certain integral equations of the first kind. *J ACM*. 1962;9(1):84-97.

doi: 10.1145/321105.321114

32. Hansen PC, O'Leary DP. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J Sci Comput*. 1993;14(6):1487-1503.

doi: 10.1137/0914086

33. Castellanos JL, Gómez S, Guerra V. The triangle method for finding the corner of the L-curve. *Appl Numer Math*. 2002;43(4):359-373.

doi: 10.1016/S0168-9274(01)00179-9

34. Haney MM, Tsai VC. Perturbational and nonperturbational inversion of Rayleigh-wave velocities. *Geophysics*. 2017;82(3):F15-F28.

doi: 10.1190/geo2016-0397.1

35. Ortega-Culaciati F, Simons M, Ruiz J, Rivera L, Díaz-Salazar N. An EPIC Tikhonov regularization: Application to quasi-static fault slip inversion. *J Geophys Res Solid Earth*. 2021;126(7):e2020JB021141.

doi: 10.1029/2020JB021141

36. Okada H. Theory of efficient array observations of microtremors with special reference to the SPAC method. *Explor Geophys*. 2006;37(1):73-85.

doi: 10.1071/EG06073

37. Cheng F, Xia J, Xu Y, Xu Z, Pan Y. A new passive seismic method based on seismic interferometry and multichannel analysis of surface waves. *J Appl Geophys*. 2015;117:126-135.

    doi: 10.1016/j.jappgeo.2015.04.005

38. Wang K, Qian J, Zhang H, Gao J, Bi D, Gu N. Seismic imaging of mine tunnels by ambient noise along linear arrays. *J Appl Geophys*. 2022;203:104718.

    doi: 10.1016/j.jappgeo.2022.104718

39. Langston CA. Spatial gradient analysis for linear seismic arrays. *Bull Seismol Soc Am*. 2007;97(1):265-280.

    doi: 10.1785/0120060100

40. Hayashi K, Asten MW, Stephenson WJ, *et al*. Microtremor array method using spatial autocorrelation analysis of Rayleigh-wave data. *J Seismol*. 2022;26:601-627.

    doi: 10.1007/s10950-021-10051-y

41. Nishida K, Takagi R, Takeo A. Ambient noise multimode surface wave tomography. *Prog Earth Planet Sci*. 2024;11:4.

    doi: 10.1186/s40645-023-00605-8

42. Haney MM, Mikesell TD, van Wijk K, Nakahara H. Extension of the spatial autocorrelation (SPAC) method to mixed-component correlations of surface waves. *Geophys J Int*. 2012;191(1):189-206.

    doi: 10.1111/j.1365-246X.2012.05597.x

43. Hansen PC. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev*. 1992;34(4):561-580.

    doi: 10.1137/1034115

44. Yu D, Hwang C, Zhu H, Ge S. The Tikhonov-L-curve regularization method for determining the best geoid gradients from SWOT altimetry. *J Geod*. 2023;97:93.

    doi: 10.1007/s00190-023-01783-5

45. Xia J, Miller RD, Park CB. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics*. 1999;64(3):691-700.

    doi: 10.1190/1.1444578

46. Mordret A, Landès M, Shapiro NM, Singh SC, Roux P. Ambient noise surface wave tomography to determine the shallow shear velocity structure at Valhall: Depth inversion with a Neighbourhood Algorithm. *Geophys J Int*. 2014;198(3):1514-1525.

    doi: 10.1093/gji/ggu217

47. Cao B, Wang J, Du H, Tan Y, Liu G. Research on comprehensive detection and visualize of hidden cavity goaf. *Sci Rep*. 2022;12:22309.

    doi: 10.1038/s41598-022-26680-3

48. Cao A, Dou L, Cai W, Gong S, Liu S, Jing G. Case study of seismic hazard assessment in underground coal mining using passive tomography. *Int J Rock Mech Min Sci*. 2015;78:1-9.

    doi: 10.1016/j.ijrmms.2015.05.001

49. Zhou L, Jia B, Bao X, Chen H, Zheng K. Construction and experimental verification of wave velocity model for source location in goaf overlying rock strata. *J Appl Geophys*. 2024;230(2):105515.

    doi: 10.1016/j.jappgeo.2024.105515

50. Arifuggaman A, Zhang C, Feng M, Chen Y, Li Q, Wang T. Mining-induced subsidence predicting and monitoring: A comprehensive review of methods and technologies. *Geotech Geol Eng*. 2025;43:314.

    doi: 10.1007/s10706-025-03271-3

51. Barton N. Some new Q-value correlations to assist in site characterisation and tunnel design. *Int J Rock Mech Min Sci*. 2002;39(2):185-216.

    doi: 10.1016/S1365-1609(02)00011-4

52. Ghabraie B, Ren G, Zhang X, Smith J. Physical modelling of subsidence from sequential extraction of partially overlapping longwall panels. *Int J Coal Geol*. 2015; 140:71-83.

    doi: 10.1016/j.coal.2015.01.004

53. Bell FG, Stacey TR, Genske DD. Mining subsidence and its effect on the environment: Some differing examples. *Environ Geol*. 2000;40(1-2):135-152.

    doi: 10.1007/s002540000140

54. Martínez-Pagán P, Navarro M, Pérez-Cuevas J, Alcalá FJ, García-Jerez A, Vidal F. Shear-wave velocity based seismic microzonation of Lorca city (SE Spain) from MASW analysis. *Near Surf Geophys*. 2014;12(6):412-423.

    doi: 10.3997/1873-0604.2014032

55. Leucci G, De Giorgi L. Microgravimetric and ground penetrating radar geophysical methods to map the shallow karst network in a coastal area (Marina di Capilungo, Lecce - Italy). *Explor Geophys*. 2010;41(3):178-188.

    doi: 10.1071/EG09029

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

GMLAN: Grouped-residual and multi-scale large-kernel attention network for seismic image super-resolution

Anxin Zhang[1†], **Zhenbo Guo**[2†*], **Shiqi Dong**[1†], **and Zhiqi Wei**[2]

[1]Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology (Ministry of Education), School of Electrical Engineering, Northeast Electric Power University, Jilin, China

[2]Bureau of Geophysical Prospecting, China National Petroleum Corporation, Zhuozhou, Hebei, China

(This article belongs to the *Special Issue: Advanced Artificial Intelligence Theories and Methods for Seismic Exploration*)

Abstract

The resolution of seismic images significantly impacts the accuracy of subsequent seismic interpretation and reservoir location. However, the resolution of seismic images often degrades due to the influence of multiple factors, making super-resolution of seismic images essential and critical. We propose a grouped-residual and multi-scale large-kernel attention network (GMLAN) framework, trained on synthetic seismic images to achieve excellent seismic image super-resolution on field seismic data. GMLAN is primarily composed of two modules: The feature extraction module (FEM) and the image reconstruction module (IRM). The FEM consists of two components: Shallow feature extraction (SFE) and deep feature extraction (DFE). The SFE component is designed to capture the basic information of seismic images, such as large-scale structures and morphological features of the strata. The DFE component serves as the cornerstone of the feature extraction process, leveraging residual groups and multi-scale large-kernel attention to distill detailed features from seismic images, such as stratigraphic interfaces, dip angles, and relative amplitudes. Finally, the IRM utilizes sub-pixel convolution, a learnable upsampling technique, to reconstruct super-resolution seismic images while preserving the continuity of seismic features. The framework demonstrates satisfactory performance on both synthetic and field data.

Keywords: Seismic images; Super-resolution; Deep learning; Grouped-residual structures; Malti-scale large-kernel self-attention

1. Introduction

Geological interpretation is highly dependent on the quality of seismic images. However, seismic images, which are obtained after data acquisition, processing, and imaging, are inevitably influenced by the acquisition environment and data processing methods, resulting in blurred events and noise. Low-resolution seismic images are detrimental to subsequent geological interpretation (e.g., fault detection and reservoir location) and

---

[†] *These authors contributed equally to this work.*

***Corresponding author:**
Zhenbo Guo
(guozhenbo01@cnpc.com.cn)

Citation: Zhang A, Guo Z, Dong S, Wei Z. GMLAN: Grouped-residual and multi-scale large-kernel attention network for seismic image super-resolution. *J Seismic Explor*. 2025;34(5):36-52.
doi: 10.36922/JSE025350063

Received: August 25, 2025

Revised: October 17, 2025

Accepted: October 23, 2025

lead to the loss of valuable details (e.g., folds, faults, and thin layers). Consequently, enhancing the resolution of seismic images has become a critical step in the seismic exploration process.

Over the past few decades, researchers have conducted numerous studies to improve the resolution of seismic images. In general, methods for enhancing resolution can be divided into two categories: Data acquisition[1-6] and data processing.[7] Data acquisition mainly includes broadband acquisition[1-3] and high-density acquisition.[4,5] Broadband acquisition increases the longitudinal resolution of seismic data by recording data over a wider frequency range, while high-density acquisition increases the number of sources and receivers and reduces the processing unit area, effectively suppressing noise and improving the resolution of seismic profiles. However, both broadband acquisition and high-density acquisition require high costs during the data acquisition process. While the data processing methods include, but are not limited to, inversion, denoising, interpolation, attenuation compensation, and deconvolution,[7,8] traditional resolution enhancement methods for seismic images typically involve multiple steps, each introducing corresponding errors.[9] The cumulative effect of these errors across multiple stages often prevents the full restoration of seismic image details in the final high-resolution output. In contrast, end-to-end deep learning approaches can perform super-resolution in a single step, thereby effectively removing the error accumulation inherent in traditional methods. Moreover, deep learning models can automatically learn and update their learnable parameters[10,11] and generally require lower computational resources and costs compared to traditional data processing techniques.

In recent years, traditional resolution enhancement methods for seismic images have become insufficient to meet the demands for rapid and high-precision processing of large amounts of seismic data. With the rapid advancement of graphics processing units (GPUs) and deep learning algorithms, deep learning methods have been widely applied in direct end-to-end processing of seismic images, enabling resolution enhancement in a single step. Among existing approaches, the convolutional neural network (CNN) is a classical method used for image super-resolution reconstruction, extracting local features of seismic images through multiple convolution operations.[12,13] During CNN training, the model learns the texture and edge information of seismic images, thereby restoring useful image content and improving resolution. Li *et al.*[14] proposed using a U-Net to achieve seismic image super-resolution. The encoder of U-Net captures contextual information through convolutional layers, pooling layers, and down-sampling operations, while the decoder gradually restores image features and improves resolution through upsampling and convolution operations, aided by skip connections. Min *et al.*[15] introduced D2Unet, a dual-decoder network based on the U-Net architecture, which simultaneously addresses the restoration of high-resolution image reconstruction and edge detection, thereby enhancing resolution while preserving critical edge information. In recent years, transformer-based networks have also achieved great success in improving the resolution of seismic images.[16,17] The self-attention mechanism, a cornerstone of the transformer architecture, dynamically attends to contextual information across relevant elements within a sequence, effectively capturing long-range dependencies and enhancing both model efficiency and expressiveness through parallel computation. The multi-head attention mechanism further enhances the model's feature extraction capacity by enabling simultaneous learning of input characteristics from multiple perspectives. In addition, in the field of seismic image super-resolution, other notable techniques include generative adversarial networks[18,19] and diffusion models.[20]

In this study, inspired by the CNN and the self-attention mechanism in the transformer, the authors propose a grouped-residual and multi-scale large-kernel attention network (GMLAN) to achieve seismic image super-resolution, combining convolution and multi-scale attention mechanisms. GMLAN is mainly composed of a feature extraction module (FEM) and an image reconstruction module (IRM). The FEM includes shallow feature extraction (SFE) and deep feature extraction (DFE). SFE utilizes convolutional layers to extract prominent large-scale features in seismic images, such as geological structures and shapes. DFE integrates group residuals and multi-scale large-kernel self-attention mechanisms to efficiently extract multi-scale features, reducing computational parameters while capturing detailed information such as layer interfaces and dip angles. The features extracted by SFE and DFE significantly influence the accuracy of subsequent image reconstruction. In the IRM, the fused shallow and deep features are upsampled and reconstructed through sub-pixel convolution to generate the final high-resolution seismic image. The authors construct a synthetic dataset whose features resemble those of field seismic images using a convolutional model and conduct supervised training on the GMLAN network. Consequently, the proposed method serves as a super-resolution processing approach for seismic images that effectively integrates algorithmic design with data. GMLAN provides a structural foundation for multi-scale feature capture and high-frequency recovery, achieved through its grouped-residual structures and multi-scale large-kernel attention (MLKA) mechanisms. Meanwhile, the synthetic dataset, which is consistent with field data in geological and seismic characteristics, provides sufficient supervised training samples.[21] The effectiveness of the proposed model is validated through tests on both synthetic and field seismic images, demonstrating superior accuracy in recovering folds and faults in field seismic data.

In the comparative experiment section, the proposed approach is compared with U-Net, demonstrating that it not only achieves significant improvements in seismic image super-resolution but also exhibits faster convergence.

2. Methods

2.1. Network architecture

The architecture of the proposed GMLAN is shown in Figure 1. It combines convolution with multiple attention mechanisms and mainly consists of two parts: FEM and IRM. The FEM can be further subdivided into SFE and DFE.

The network first performed SFE on the input low-resolution seismic images, transforming the image space into a higher-dimensional feature space. Subsequently, the DFE was employed to establish a non-linear mapping relationship between low-resolution and high-resolution images, thereby reconstructing higher-frequency texture details. After feature extraction, residual connections were introduced to integrate the extracted shallow features with the deep features. This approach not only facilitates the extraction of high-dimensional image information but also ensures amplitude preservation in super-resolution reconstruction and reduces artifacts. In the final IRM, the fused features were first integrated, and their channels were adjusted using convolutional layers. Subsequently, the integrated features were upsampled via sub-pixel convolution to enhance the edge and texture details of the seismic image, thereby improving the resolution and yielding the final super-resolution reconstructed images.

The proposed GMLAN was trained on synthetic data, and the well-trained GMLAN was applied to the super-resolution prediction of field seismic images.

2.1.1. Feature extraction

The SFE was primarily implemented by a convolutional layer with a 3×3 convolution kernel. It can be expressed by the formula:

$$X_0 = L_{SF}\left(I_{in}\right) \tag{I}$$

Where $I_{in} \in R^{H\times W\times C_{in}}$ and $X_0 \in R^{H\times W\times C_s}$ refer to the low-resolution input image of the network and the output feature map of SFE, respectively. H and W represent the height and width of the input seismic images and feature maps, respectively, C_{in} and C_s denote the number of channels in the input image and the output feature maps of the shallow features, respectively. L_{SF} refers to the main layer of SFE.

In the DFE part, multi-branch residuals and large-kernel self-attention are the core components, which were executed in four stages. In DFE, four grouped residual and MLKA (GRMLKA) blocks were employed, corresponding to the four stages of DFE. In each stage of DFE, six layers of multi-residual groups were set. Furthermore, MLKA mechanisms were incorporated after the multi-branch residuals in the second, fourth, and sixth layers to enhance the precision and accuracy of feature extraction. Therefore, each DFE stage can actually be divided into three subgroups. It is worth noting that the number of multi-branch residual groups in each DFE stage should be set to an even number because the MLKA module was only added to the even-numbered layers. The operation sequence of each stage is as follows: first, pass through the first multi-branch residual block; then, perform normalization and pass through the feed-forward network (FFN); finally, conduct another normalization operation to complete the first layer of operations. In the second layer, unlike the first layer, a MLKA operation was carried out after the first multi-branch residual group. After this operation, normalization and subsequent processes were performed. This set of operations can be expressed by the following formula:

$$\begin{aligned}
X_{i,1} &= L_{RG}\left(X_{i-1}\right)\\
X_{i,2} &= L_{Norm}\left(X_{i,1}\right) + X_{i-1}\\
X_{i,3} &= L_{FFN}\left(X_{i,2}\right)\\
X_{i,4} &= L_{Norm}\left(X_{i,3}\right) + X_{i,2}\\
X_{i,5} &= L_{RSA-MLKAB}\left(X_{i,4}\right)\\
X_{i,6} &= L_{Norm}\left(X_{i,5}\right) + X_{i,4}\\
X_{i,7} &= L_{FFN}\left(X_{i,6}\right)\\
X_{i,8} &= L_{Norm}\left(X_{i,7}\right) + X_{i,6}
\end{aligned} \tag{II}$$

The entire process of extracting deep features is expressed as:

$$\begin{aligned}
X_i &= L_{GRMLKA}\left(X_{i-1}\right)\\
I_{DF} &= L_{Conv}\left(X_N\right) + X_0
\end{aligned} \tag{III}$$

Where X_i, X_{i-1} refer to the output of each GRMLKA, and the range of i is $1 \le i \le N$. Given that four stages were set in our model, we set $N = 4$. I_{DF} denotes the final output feature image obtained from DFE. L_{GRMLKA} represents the

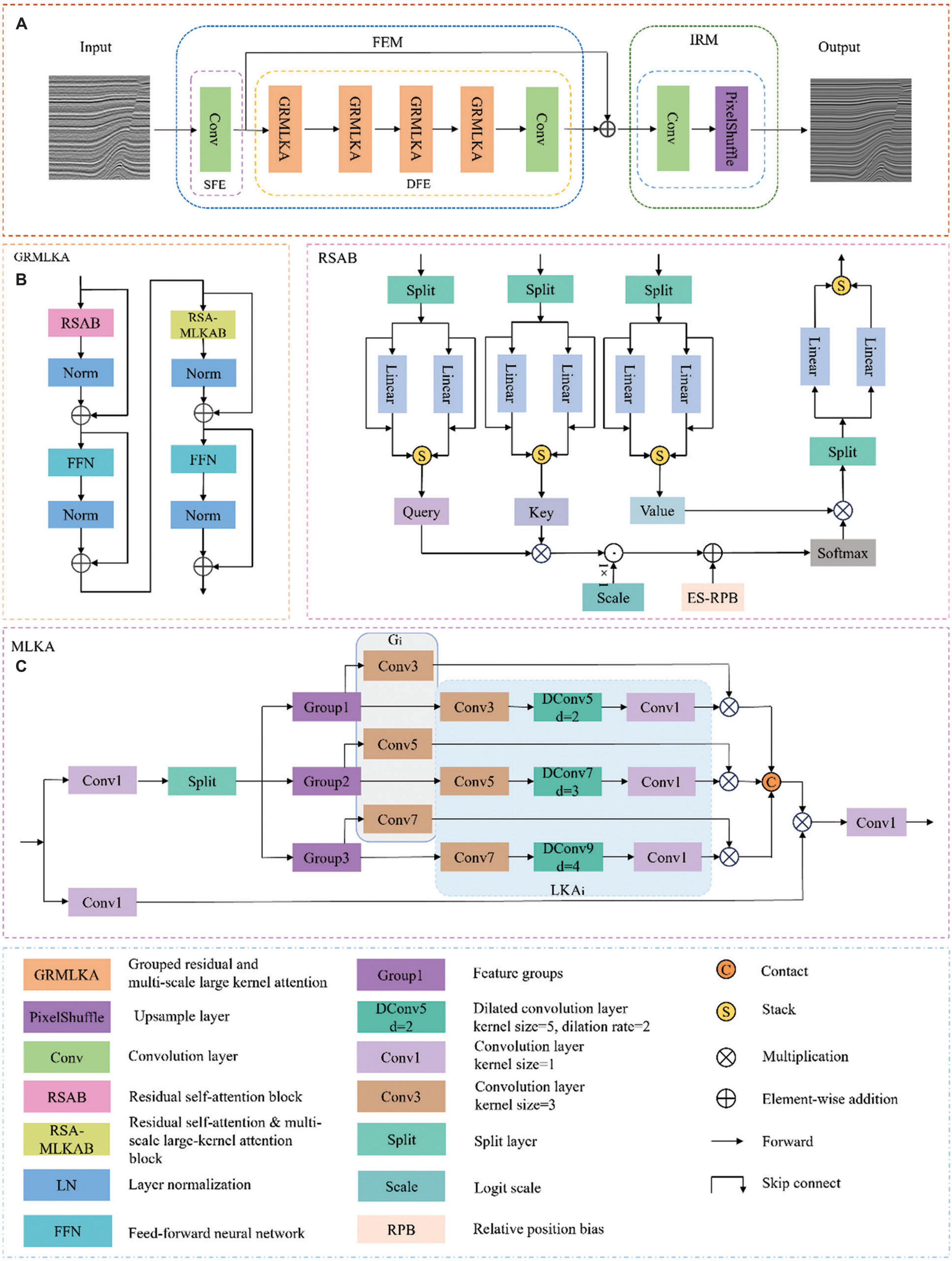


Figure 1. The architecture and components of the grouped-residual and multi-scale large-kernel attention network

four stages within the DFE process, and L_{Conv} represents the last convolutional layer within the DFE. Within each GRMLKA stage, the output feature map of each block is denoted as $X_{i,m}$. In **Equation (II)**, the index *m* represents integers from 1 to 8. Given that each stage is divided into three subgroups, the range of *m* is $1 \le m \le 24$. Specifically, when $m = 24$, $X_{i,24}$ is simply denoted as X_i, representing the output of the *i-th* GRMLKA. L_{RG} refers to the residual group layer, L_{Norm} denotes the layer normalization component, L_{FFN} indicates the FFN layer, and $L_{RSA\text{-}MLKAB}$ represents the fusion layer integrating residual groups and large-kernel attention.

The self-attention mechanism is highly effective at capturing long-range spatial dependencies within seismic images.[22] However, the process of generating the query (*Q*), key (*K*), and value (*V*) matrices and calculating their products often leads to computational redundancy.[23] To harness the benefits of the self-attention mechanism while mitigating this redundancy, we incorporated a residual self-attention block (RSAB) into the proposed GMLAN model, as depicted in Figure 1B. Specifically, we partitioned the input feature channels into two distinct groups and applied separate linear transformations and residual calculations to each group. The following equation illustrates the operation of grouping the input feature vectors:

$$\left[A_1, A_2\right] = S\left(A\right) \tag{IV}$$

Where *A* refers to the input feature vector. The operation *S* pertains to the grouping of features by channels, which are divided into two distinct groups, denoted as A_1 and A_2, with each group encompassing half of the channels.

We then solved each component and incorporated residual connections, merging the outcomes from the two branches to derive the *Q*, *K*, and *V* matrices. To illustrate, the computation of the *Q* matrix is detailed in the following formula:

$$Q^{'} = ST\left(A_1 + Linear\left(A_1\right), A_2 + Linear\left(A_2\right)\right) \tag{V}$$

Where

$$Linear\left(A_1\right) = A_1 \times W_{q1}{}^{T} + b_{q1} \tag{VI}$$

$$Linear\left(A_2\right) = A_2 \times W_{q2}{}^{T} + b_{q2} \tag{VII}$$

ST refers to the operation of stacking along a specified dimension. The term $Q^{'} \in R^{2\times B\times N\times C/2}$ represents the outcome after the merging operation. *B* signifies the batch size, *N* indicates the number of tokens obtained after flattening the input features, and *C* represents feature channels. *Linear* denotes the application of linear mapping operations to A_1 and A_2 individually. $W_{q1}{}^{T}$ and $W_{q2}{}^{T}$ are weight matrices, while b_{q1} and b_{q2} denote relative position bias terms. Following the computation of $Q^{'}$, feature rearrangement and merging operations were conducted to obtain the final *Q* matrix. The *K* and *V* matrices were derived using an analogous approach and were subsequently normalized. The attention was then calculated using the formula provided below:

$$Attention\left(Q,K,V\right) = Softmax\left(\left(Q \times K^{T}\right) \times Scale + B\right) \times V \tag{VIII}$$

Where *Q*, *K*, and *V* were normalized, *Scale* is a learnable scaling factor that dynamically adjusts the distribution of attention scores, enabling the model to adaptively learn the optimal similarity scale and enhance its focus on significant features. *B* refers to the continuous relative position bias generated by a multi-layer perceptron, which can effectively accommodate the long-range structural continuity inherent in seismic data. When combined with a grouped residual design, it enabled independent optimization of long-range association computations within each feature group. The attention mechanism implemented in RSAB is similar to cosine attention but differs using a learnable scaling factor instead of a fixed value and by incorporating continuous relative position biases. The resulting attention scores were individually subjected to linear projections and then combined to produce the final output of this module. This not only minimizes cross-channel interference but also halves the computational load, thereby avoiding redundancy in long-range calculations.

The grouped learning and computation operation not only captures interdependencies among different channels but also reduces computational complexity.[24] Specifically, before performing grouped calculations, each group has *C* channels, and the computational cost is C^2. After grouping, both the number of channels and the computational cost are halved compared to their values before grouping.

In the DFE stage, in addition to the grouped residual attention, we introduced a MLKA mechanism, as depicted in Figure 1C. Convolution kernels of varying sizes enabled the capture of seismic image features at multiple scales, thereby enhancing the interconnections among different receptive fields. Within the MLKA module, the output from the preceding stage was initially subjected to convolution to augment the number of input feature channels, facilitating the subsequent division of feature channels. Assuming the input features have *n* channels, the output features after this operation will have 2*n* channels, thereby increasing the non-linearity of the seismic image features. This step can be formulated as follows:

$$\left[X_{1:n}, X_{n+1:2n}\right] = C\left(X_{RSABO}\right) \tag{IX}$$

Where X_{RSABO} represents the output features from the RSAB. The operation C denotes the process of augmenting the feature channels and partitioning them into two identical groups. The notations $X_{1:n}$ and $X_{n+1:2n}$ refer to the two segments of the partitioned features, where $1{:}n$ indicates the range of the first half of the feature channels after dimensionality increase, and $X_{1:n}$ can alternatively be denoted as X_a. Similarly, $X_{n+1:2n}$ signifies the range of the latter half of the feature channels, and thus is also referred to as X_b. Both X_a and X_b have the same number of feature channels as X_{RSABO}.

Consequently, the input features, following the augmentation of feature channels, were split into two equivalent segments. Each segment underwent a convolutional layer to adjust the number of output feature channels, aligning them with the input requirements of subsequent operations. One segment was further subdivided into three parts along the feature channels for the upcoming large kernel attention ($\mathrm{LKA_i}$) operation. These steps can be mathematically formulated as follows:

$$F_{Gn} = S\left(L_{C1}\left(X_a\right)\right) \tag{X}$$

Here, L_{C1} denotes the convolutional layer with a kernel size of 1 in Figure 1C, while F_{Gn} signifies different feature groups, with the subscript n indicating the group identifier, which can extend up to 3. Within the $\mathrm{LKA_i}$ module, there were three analogous pathways that processed the corresponding feature groups independently, namely Group1 to Group3 as depicted in the figure. Each pathway's LKA comprised three distinct convolutional layers. The initial layer consisted of standard convolutional layers with kernel sizes of 3 × 3, 5 × 5, and 7 × 7, enabling the extraction of local image features at varying scales. The smaller the convolution kernel, the more detailed the geological structures captured in the extracted feature map, indicating a richer content of high-frequency information. The subsequent layer employed dilated convolution with different kernel sizes and dilation rates; this differs from the first layer by expanding the receptive field of the convolutional kernel, thereby allowing it to capture broader global features.[25] By integrating small-kernel convolutions with dilated convolutions, the model can effectively capture both low-frequency information, such as large-scale geological structures in seismic data, and high-frequency features. This integration compensates for the insufficiencies in learning high-frequency components that can occur with large-kernel convolutions alone.[26] The final convolutional layer was utilized to modulate the feature channels and enhance the non-linearity of the output features. The procedure of LKA can be mathematically represented as follows:

$$LKA_i\left(F_{Gn}\right) = L_{C1}\left(L_{DC}\left(L_C\left(F_{Gn}\right)\right)\right) \tag{XI}$$

Where L_C refers to the initial standard convolutional layer within the LKA_i module, and L_{DC} signifies the dilated convolutional layer. However, convolution with large kernels may introduce artifacts. To address this issue, we incorporated spatial gating (G_i) into the network to enhance the image's feature representation capabilities.[27] The configuration of G_i mirrored that of the initial convolutional layer in the respective branches of $\mathrm{LKA_i}$. The convolutional layers with different kernel sizes in the spatial gating focused on the weak signal features of varying intensities in the seismic images. After passing through G_i, the grouped features were integrated with the output features from $\mathrm{LKA_i}$. The resulting three sets of integrated features were concatenated along the channel dimension to yield a feature map X_a', which matches X_a in terms of channel count. The final output of the MLKA module can also be viewed as an attention computation, which is mathematically expressed as follows:

$$O_{MLKA} = L_{C1}\left(CAT\left(LKA_i\left(F_{Gn}\right) \times G_i\left(F_{Gn}\right)\right)\right) \tag{XII}$$

Where CAT denotes the operation of concatenating along the feature channels, and O_{MLKA} refers to the overall output of the MLKA module.

The application of the MLKA module within the DFE not only facilitated the extraction of local features of seismic images at various scales but also leveraged dilated convolutions to expand the receptive field without additional computational overhead. This approach effectively addresses the issue of missing long-range feature mappings that can arise from focusing solely on local feature extraction.

2.1.2. Image reconstruction

In the IRM, we employed a 3 × 3 convolutional layer to adjust the channels of the fused deep and shallow features. Finally, through the upsampling layer, an image with dimensions twice those of the input image in each direction was generated, effectively reconstructing the image to a super-resolution version $I_{SR} \in R^{H \times W \times C_{out}}$.

I_{SR} refers to the reconstructed seismic super-resolution image, which also serves as the final output of the entire network, while H, W, C_{out} denote the height, width, and number of channels of the output image, respectively.

The entire image reconstruction process can be represented by the following formula:

$$I_{SR} = L_{upsample}\left(L_{Conv}\left(I_{DF} + I_{SF}\right)\right) \tag{XIII}$$

Where I_{SR} denotes the reconstructed super-resolution image, I_{SF} refers to the extracted shallow features, and I_{DF} represents the features extracted by the DFE module. $L_{upsample}$ refers to the upsampling layer in the reconstruction part, where we employed sub-pixel convolution for upsampling. Traditional interpolation techniques, such as bilinear and nearest neighbor interpolation, can enlarge low-resolution images to high-resolution counterparts. However, these methods often introduce image artifacts and blurring. In contrast, sub-pixel convolution, which utilizes convolutional operations and channel recombination, can enlarge low-resolution images without distortion. Furthermore, it significantly enhances image clarity and detail, thereby preserving a greater amount of detailed information.

2.2. Modeling of the synthetic dataset

For the training of GMLAN, we used synthetic seismic images. The generation of the synthetic dataset follows the method proposed by Wu *et al.*[28] A total of 2,000 pairs of synthetic data were generated, with 1,600 pairs allocated for training and the remaining 400 pairs reserved for validation and testing. Given that the GMLAN employed supervised training, it required a substantial amount of paired low-resolution and high-resolution image data. The high-resolution images served as the ground truth, while the low-resolution images were used as input. However, the availability of high-resolution field seismic images was limited, which restricted the amount of data that could serve as ground truth. In addition, the cost of obtaining high-resolution data was substantial. Therefore, in the proposed model training, a large number of synthetic seismic images were generated for training.

The synthetic images mainly simulated the distribution of geological structures based on the target data. First, a three-dimensional reflectivity model was established, and fold and fault structures were added to all horizontal layers. The frequency band of the high-resolution image was wider than that of the low-resolution image. Therefore, to obtain the corresponding high-resolution and low-resolution data, convolution operations were performed between the three-dimensional reflectivity models (with added folds and faults) and wavelets of different frequencies. Specifically, the three-dimensional reflectivity models were convolved with a high-frequency Ricker wavelet (35–55 Hz) to obtain the high-frequency seismic volumes, and then with a low-frequency Ricker wavelet (10–25 Hz) to obtain the corresponding low-frequency seismic volumes. Two-dimensional seismic slices were extracted from the high-frequency seismic volumes to serve as labels for actual training. Colored noise was added to the low-frequency seismic volumes to simulate real-world interference. Two-dimensional slices were then extracted from the low-frequency seismic volume using the same method and downsampled to obtain the low-resolution data required for training. The downsampled low-resolution seismic data and the high-resolution labels were paired. Finally, using the above-described method, 2,000 pairs of synthetic seismic data were generated and allocated for training, validation, and testing with an 8:1:1 ratio.

2.3. Loss function

In this study, we used a combined loss function consisting of the L1 loss and the multi-scale structural similarity index measure (MS-SSIM):

$$L_{Mix} = a \times L_{MS-SSIM} + \left(1-a\right) \times L_1 \tag{XIV}$$

Where

$$L_{MS-SSIM} = 1 - MS - SSIM(x,y) \tag{XV}$$

L_{Mix} refers to the combined loss function, L_1 denotes the L1 loss function, and $L_{MS\text{-}SSIM}$ signifies the MS-SSIM loss. The parameter a represents the weight of the $L_{MS\text{-}SSIM}$ component. As the sum of the weights for L_1 and $L_{MS\text{-}SSIM}$ equals 1, the weight assigned to L_1 was set to $1-a$. $MS\text{-}SSIM(x, y)$ denotes the MS-SSIM, which was used to assess the similarity between the target image x and the output image y produced by the GMLAN across multiple scales. The L1 loss function is less sensitive to outliers, thereby demonstrating greater robustness in image reconstruction tasks. Meanwhile, MS-SSIM focuses on the structural information and perceptual quality of seismic images, evaluating image similarity across multiple scales. This multi-scale evaluation enables the model to capture richer details and enhance performance across different scales.[29]

Multi-scale structural similarity is an improvement over the structural similarity index measure (SSIM), which evaluates the SSIM at different scales and aggregates these values to obtain MS-SSIM:

$$MS - SSIM\left(x,y\right) = \left[l_M\left(x,y\right)\right]^{\alpha_M} \times \prod_{j=1}^{M}\left[c_j\left(x,y\right)\right]^{\beta_j} \times \left[s_j\left(x,y\right)\right]^{\gamma_j} \tag{XVI}$$

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)} \tag{XVII}$$

Where M represents the number of scales (set to 5 in this study). α_M is the weight of the fifth scale, which equals 0.1333; α, β, and γ are the weights vector for each part, set to $\alpha = \beta = \gamma$ [0.0448, 0.2856, 0.3001, 0.2363, 0.1333], x and y are the output image of GMLAN and the label image, respectively; μ, σ^2, σ_{xy} represent the mean, variance, and covariance of the corresponding images. C_1 and C_2 are stabilizing terms used to prevent division by zero. Structural similarity takes into account the luminance $l(x,y)$, contrast $c(x,y)$, and structural $s(x,y)$ information of the image, providing a more comprehensive evaluation that aligns more closely with the human visual system's perception of image quality.

Among them, brightness, contrast, and structural similarity can be expressed as:

$$l(x,y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1} \tag{XVIII}$$

$$c(x,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2} \tag{XIX}$$

$$s(x,y)=\frac{\sigma_{xy}+C_2/2}{\sigma_x\sigma_y+C_2/2} \tag{XX}$$

After pre-training tests, we set $a = 0.6$.

2.4. Evaluation metrics

The peak signal-to-noise ratio (PSNR) is an indicator used to measure the quality of image reconstruction. It calculates the pixel-level error between the ground truth and the network's output based on the mean squared error (MSE). PSNR provides an intuitive numerical value describing the accuracy of the reconstructed image. Generally, the larger the PSNR value, the higher the quality of the reconstructed image. The PSNR can be expressed as follows:

$$PSNR=10\times\log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{XXI}$$

Where

$$MSE=\frac{1}{\text{n}}\sum_{i=1}^{n}\left(y_i-x_i\right)^2 \tag{XXII}$$

Where MAX denotes the maximum value of the image data (set to 1), MSE is the MSE, n denotes the sample size (i.e., the number of seismic images used for training). y_i represents GMLAN's prediction for the i-th seismic image, whereas x_i denotes the true value of the i-th seismic image. However, PSNR only describes the pixel-level error between images and cannot fully reflect the perception of the human visual system, especially in terms of image details and textures.

Therefore, to address this limitation, the SSIM was introduced as another evaluation indicator during training.

The range of SSIM value is [−1, 1]; a value closer to 1 indicates a higher similarity between the super-resolution reconstructed image and the labeled image.

2.5. Experimental setup

The experiment was conducted using PyCharm (PyCharm Community Edition, version 2024.3, JetBrains, Czech Republic) and the PyTorch (version 2.5.1, Meta Platforms, Inc., United States) GPU framework. A Compute Unified Device Architecture parallel computing framework was implemented (version 12.4, NVIDIA Corporation, United States), and training was performed on two NVIDIA RTX 4090 GPUs, each with 24 GB of memory.

To enhance training efficiency, the data were normalized during training. The input data were normalized to the range of [−1, 1] using the following method:

$$x^*=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{XXIII}$$

Where x denotes the variable to be normalized, which, in the context of the experiment, corresponds to the input training data. $x_{\max}$ and $x_{\min}$ represent the maximum and minimum values of x, respectively, and x^* indicates the normalized result.

To ensure the accuracy of the predicted images, the GMLAN outputs were denormalized before calculating the evaluation metrics:

$$x^{'}=x^*\times\left(x_{\max}-x_{\min}\right)+x_{\min} \tag{XXIV}$$

Where x' signifies the denormalized output.

The parameter configurations for the experiment are detailed in Table 1.

3. Numerical tests

3.1. Comparison method

In this section, we compare the super-resolution performance of the proposed method with that of the U-Net. U-Net is highly flexible and can be adapted to various super-resolution tasks and datasets by

Table 1. Training parameters of different methods

Hyperparameter	U-Net	GMLAN (the proposed model)
Optimizer	Adam	Adam
Batch size	8	8
Epoch	400	400
Learning rate	[^{1}arn^{-3}, 1rni^{-4}, 1rni^{-5}]	[^{11}rn^{-3}, 1rni^{-4}, 1rni^{-5}]
Epochs for learning rate decay	[200, 300]	[200, 300]
Input channels	1	1
Parameters counts	17,409,537	3,377,140

Abbreviations: Adam: Adaptive moment estimation; GMLAN: Grouped-residual and multi-scale large-kernel attention network.

adjusting parameters such as network depth, width, and convolution kernel size. It has been widely applied in numerous image processing tasks, including image super-resolution reconstruction, where it has demonstrated high effectiveness and adaptability.

To ensure the fairness of the comparison experiment and the persuasiveness of the results, we trained a U-Net on the same dataset and kept the hyperparameters (such as the number of epochs, learning rate, optimizer, etc.) consistent with those used in the proposed model. The specific parameter settings are shown in Table 1. After completing parameter and data preparation, we trained GMLAN and U-Net for the same number of epochs. The comparison of the training process is shown in Figure 2.

As depicted in Figure 2A, the loss curve of the U-Net model exhibits a rapid decline during the initial 50 epochs, after which the rate of decrease gradually slows. After 200 epochs, the loss reached a stable state. In contrast, the proposed approach demonstrated a significantly faster reduction in loss at the start of training, and the loss value consistently remained lower than that of the U-Net model throughout the entire training process.

As illustrated in Figure 2B and C, the proposed method consistently outperforms U-Net in terms of both PSNR and SSIM, particularly in the later stages of training. The SSIM value of U-Net increased rapidly during the initial 50 epochs, then slowed, and ultimately stabilized at approximately 0.85. In contrast, the proposed model maintained a stable SSIM value of around 0.91. Regarding PSNR, the U-Net's value eventually stabilized at 19 dB, whereas the proposed approach achieved a value exceeding 21 dB. Therefore, considering the loss function and these two metrics comprehensively, the proposed method demonstrated significant superiority over U-Net in all training indicators.

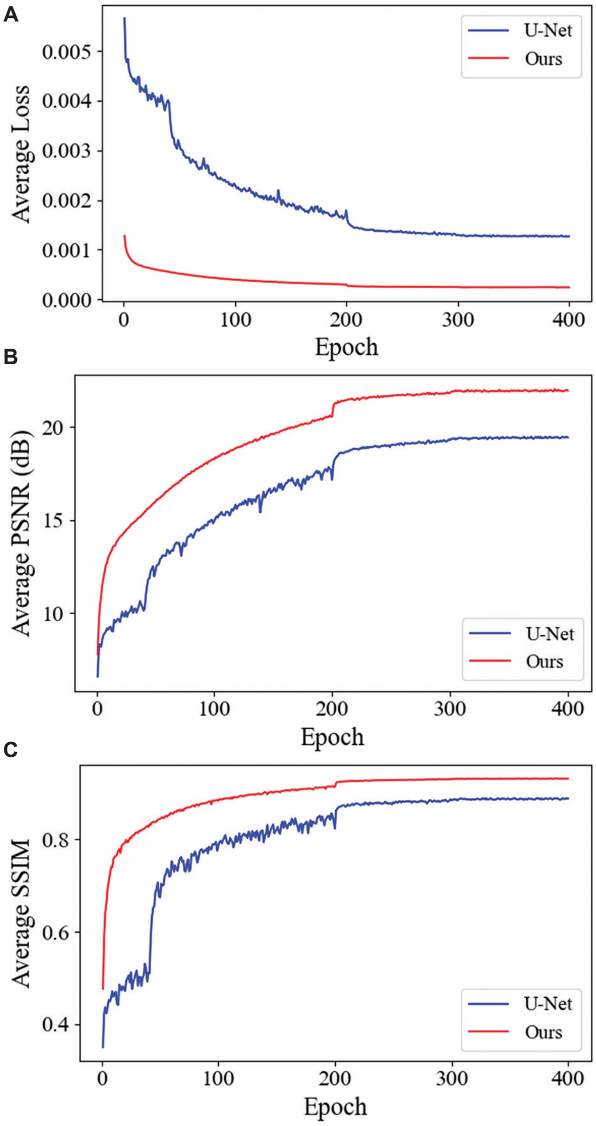


Figure 2. Comparison plot of average loss (A), peak signal-to-noise ratio (PSNR; B), and structural similarity index measure (SSIM; C) achieved by different methods with identical hyperparameters

3.2. Synthetic data testing

First, we evaluated the performance of GMLAN on the test dataset. As shown in Figure 3, at the positions indicated by the red arrows, the proposed model outperforms U-Net in fault recovery. The fault edges recovered by the proposed approach are clear and easily observable, while the fault edges within the solid red box recovered by U-Net appear smooth and blurred, resulting in the loss of edge features. In the area within the dashed yellow box, the interfaces between different layers are blurred in the prediction produced by the U-Net. We extracted the 40th trace from the data in Figure 3B-D to more effectively demonstrate the signal restoration performance. The amplitudes derived from both the proposed method and the U-Net closely resemble those of the ground truth. Nevertheless, the curve produced by the proposed model aligns more precisely with the ground truth, suggesting that the seismic images reconstructed by the proposed approach contain more detailed amplitude information. The coherence attribute serves as a highly effective and extensively utilized tool in the identification of structural and stratigraphic anomalies within seismic data.[30,31] To effectively highlight the variations in strata and the fault prediction results obtained by different methods, we have incorporated a coherence analysis, as depicted in Figure 4. The red solid-line frame within the image distinctly illustrates that the proposed approach outperforms Unet in fault recovery. This visual representation serves to underscore

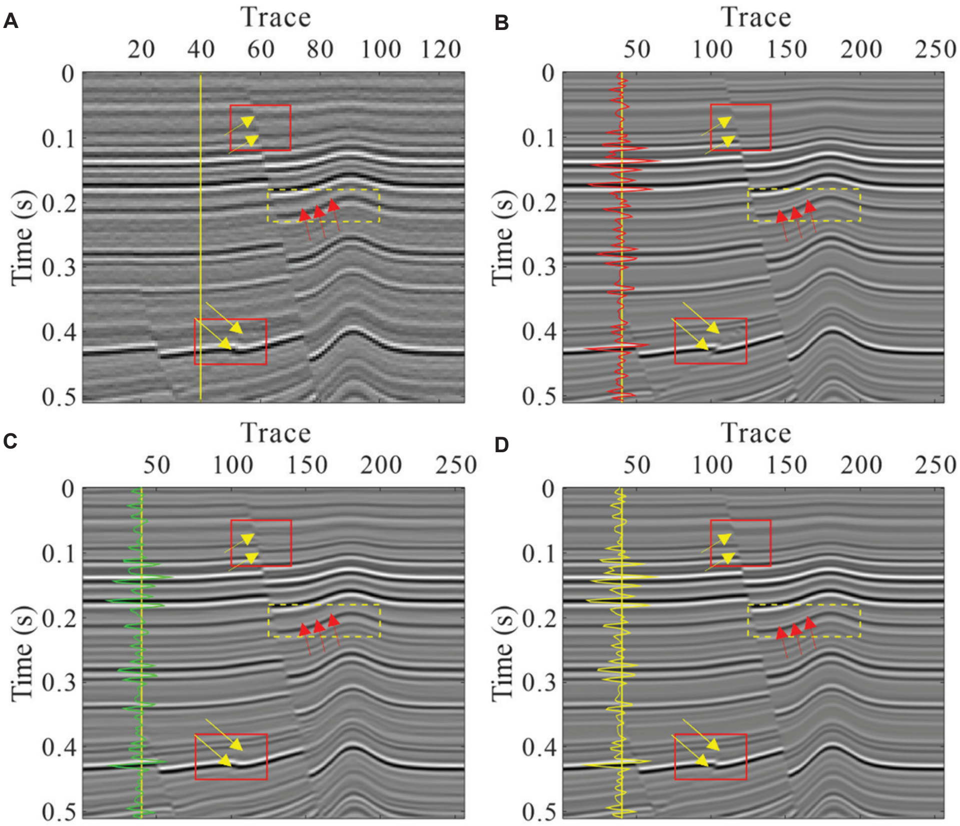


Figure 3. Comparison of super-resolution reconstruction results on synthetic data for different networks. (A) The low-resolution seismic image. (B) The ground truth of one synthetic dataset. (C) The super-resolution result predicted by the U-Net. (D) The super-resolution results predicted by the grouped-residual and multi-scale large-kernel attention network. The yellow line represents the 40th trace. In panels B, C, and D, the red, green, and yellow curves denote the amplitude of the corresponding seismic images at the 40th trace, respectively.

the enhanced capability of the proposed approach in accurately delineating geological features, which is a critical aspect of seismic image analysis. Figure 5 presents the spectra corresponding to the seismic images shown in Figure 3. Across most frequency bands, the proposed method closely matches the spectral curves of the ground truth, particularly in the mid-to-high frequency range (50–150 Hz). This consistency indicates that the proposed model excels at recovering the frequency components of low-resolution seismic images. The U-Net also approximates the ground truth with reasonable accuracy across some frequencies; however, it shows discrepancies at specific frequency points, especially in the high-frequency band. These discrepancies suggest that the U-Net is less effective than the proposed approach in recovering high-frequency components. In addition, Table 2 provides the evaluation metrics for the super-resolution prediction results of the synthetic data using the U-Net and the proposed method. The results demonstrate that the proposed model achieves significantly superior performance compared to U-Net in terms of PSNR, SSIM, and root mean square error. These findings corroborate the results presented in Figures 3-5.

3.3. Field data testing

To further verify the performance of GMLAN on field data, we selected a seismic image from the Netherlands F3 seismic survey for testing. The data, consisting of 512 traces contaminated by noise, were used as the input low-resolution seismic images. There is no ground truth for the field data. As shown in Figure 6, Figure 6A presents the field data from the Netherlands F3 block, Figure 6B illustrates the prediction result of the U-Net, and Figure 6C displays the prediction result of the proposed approach. Figures 6D-F are the coherence maps corresponding to Figure 6A-C, respectively. From the seismic images, it is evident that the prediction result of U-Net (Figure 6B) exhibits excessive smoothness at the fault locations indicated by the red arrows within the blue dashed box, leading to the loss of critical details. In contrast, the proposed method (Figure 6C) effectively restores the minor faults within the stratigraphic layers with high clarity. Within the two red solid boxes, the regions indicated by the red arrows show the recovered fault edges. The fault edges reconstructed by U-Net are less distinct compared to those recovered by

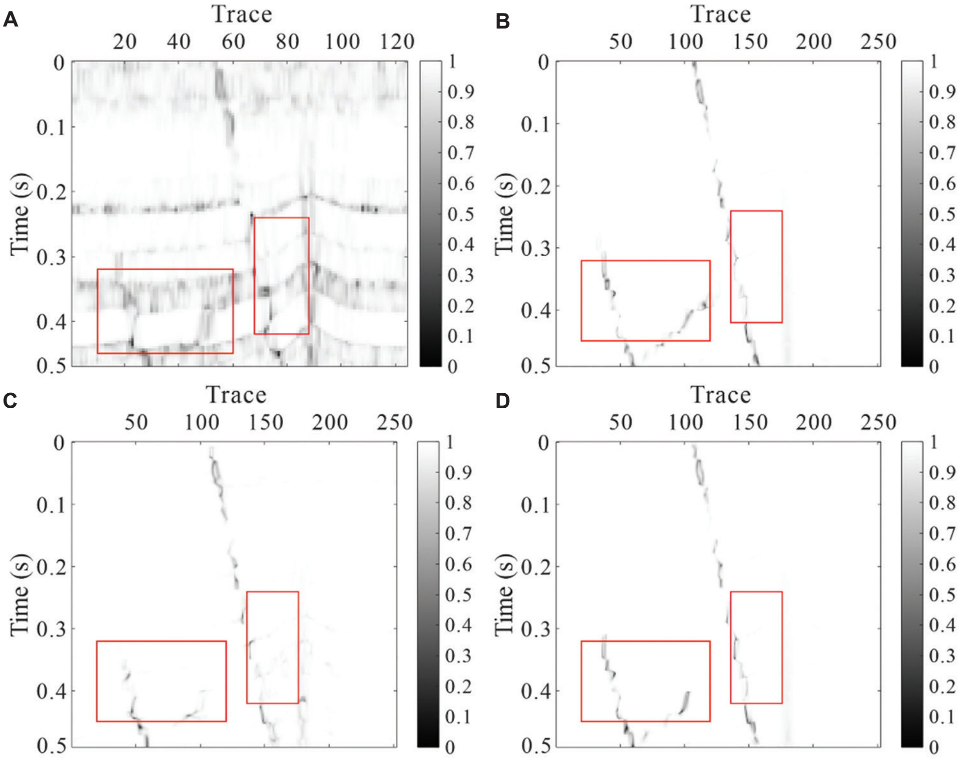


Figure 4. Comparison of seismic coherence results on synthetic data. (A) Coherence map of the low-resolution synthetic data. (B) Coherence map of the ground truth of the synthetic image. (C) Coherence map of the super-resolution result predicted by U-Net. (D) Coherence map of the super-resolution result predicted by the grouped-residual and multi-scale large-kernel attention network.

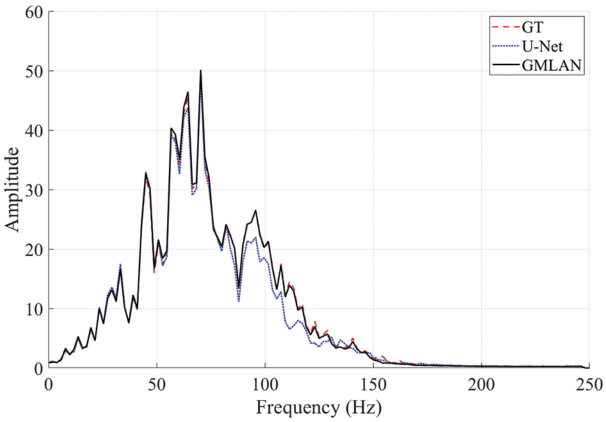


Figure 5. Spectral comparison for a synthetic seismic image
Abbreviations: GMLAN: Grouped-residual and multi-scale large-kernel attention network; GT: Ground truth.

Table 2. Comparison of evaluation metrics for super-resolution results

Method	PSNR (dB)	SSIM	RMSE
U-Net	13.68	0.68	0.25
GMLAN (the proposed model)	19.42	0.89	0.12

Abbreviations: GMLAN: Grouped-residual and multi-scale large-kernel attention network; PSNR: Peak signal-to-noise ratio; RMSE: Root mean square error; SSIM: Structural similarity index measure.

GMLAN. In Figure 6A, the position indicated by the blue arrow within the red solid boxes reveals a faintly discernible fault structure in the input low-resolution seismic image. However, as depicted in Figure 6B, the U-Net result fails to accurately restore this fault structure and instead incorrectly reconstructs it as a continuous stratigraphic layer, thereby introducing continuity artifacts into the super-resolution seismic image. In contrast, Figure 6C illustrates that GMLAN clearly restores the fault structure. This comparison with U-Net on field seismic images underscores the superior detail fidelity and structural accuracy of GMLAN. Similarly, the coherence maps in Figure 6 clearly demonstrate that, at the corresponding positions to the seismic images, the stratigraphic structure details reconstructed by the proposed model exhibit notably enhanced resolution and precision. The objective of super-resolution reconstruction for seismic images was to preserve low-frequency signals without degradation while simultaneously expanding the high-frequency bands. Figure 7 presents the average spectra of the low-resolution field seismic images and the super-resolution results reconstructed using the two different methods. In

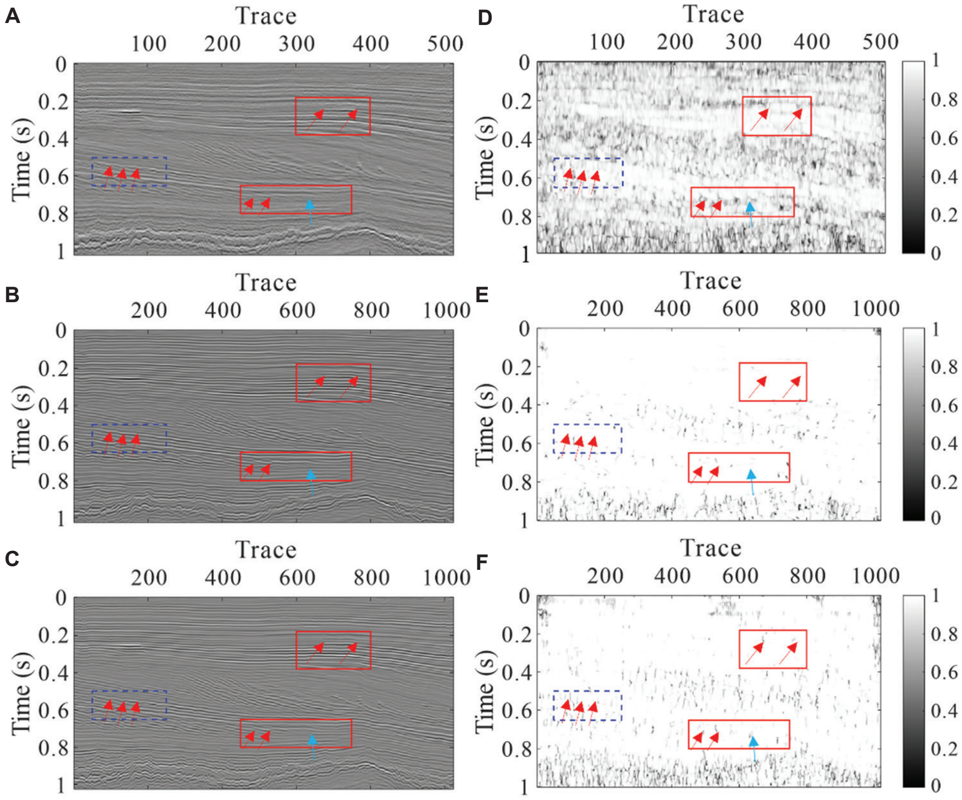


Figure 6. Comparison of super-resolution results on field data. (A) The low-resolution field data. (B) The reconstruction result of the U-Net. (C) The super-resolution reconstruction result of the proposed model. (D) Coherence map corresponding to the low-resolution field data in panel A. (E) Coherence map corresponding to the U-Net reconstruction result in panel B. (F) Coherence map corresponding to the reconstruction result of the proposed approach in panel C.

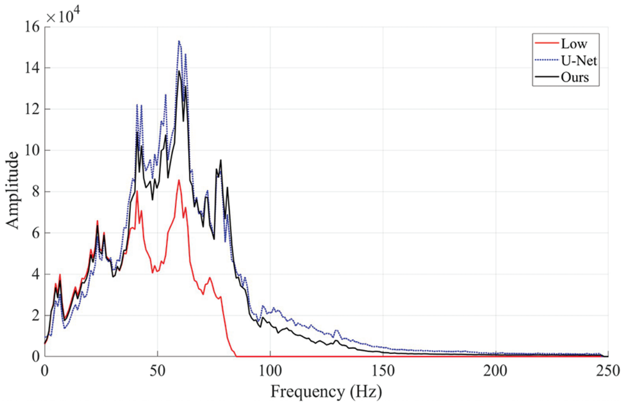


Figure 7. Comparison of average spectra of super-resolution results on field seismic images obtained by different methods.

the low-frequency range, the proposed approach exhibits a higher degree of correlation with the low-frequency components of the low-resolution seismic images. In contrast, the proposed method not only effectively expands the high-frequency band but also maintains a significant superiority in preserving the consistency of low-frequency information.

In this section, we conduct a comparative analysis between GMLAN and U-Net across various datasets. These comparisons consistently demonstrate the superiority of the proposed method in the context of seismic image super-resolution reconstruction. Furthermore, the proposed model also demonstrates robust performance on field data, highlighting its strong generalization capability and reliability when applied to diverse datasets.

4. Discussion

4.1. Ablation test

In this section, the indispensable contributions of the network components are verified through ablation experiments, including the configuration of the number of stages for the MLKA and the DFE.

First, we removed all MLKA modules from the network. This variant is referred to as the grouped residual attention network (GRAN), which was trained using the same parameter settings and training data. In addition, we conducted experiments to explore the impact of different numbers of DFE stages on network performance. The experiments in this section are summarized as follows:

(i) GRAN, without MLKA, used to verify the impact of large-kernel attention on network performance; (ii) $MLKA_2$, a network with two DFE stages; (iii) $MLKA_3$, a network with three DFE stages; (iv) $MLKA_5$, a network with five DFE stages. Given that the proposed network incorporated four DFE stages, it was designated as $MLKA_4$.

After determining the number of experiments to be conducted, we set identical hyperparameters for all four networks and trained them using the same dataset. The evaluation metrics for the training results are presented in Table 3.

From the evaluation results, it can be observed that the network without MLKA exhibited the poorest performance. Although $MLKA_5$ performed slightly better than $MLKA_4$, the improvement in PSNR and SSIM was marginal. This minor enhancement came at the cost of an additional 3,528 seconds of training time, making it computationally inefficient. In contrast, $MLKA_4$ demonstrated a slight advantage over $MLKA_3$, achieving a 1.62db increase in PSNR. This enhancement is considered reasonable given the moderate increase in training time (an additional 3,528 s). Overall, the incorporation of MLKA and the implementation of the four DFE stages prove to be the optimal configuration.

Table 3. Comparison of evaluation index results from different methods

Method	PSNR (dB)	SSIM	Training time (s)
GRAN	19.20	0.8879	70,104
$MLKA_2$	19.90	0.9080	80,160
$MLKA_3$	20.32	0.9228	83,688
$MLKA_5$	21.98	0.9324	90,744
The proposed model (GMLAN/$MLKA_4$)	21.94	0.9321	87,216

Note: The subscript attached to "MLKA" denotes the number of deep feature extraction stages.
Abbreviations: GMLAN: Grouped-residual and multi-scale large-kernel attention network; MLKA: Multi-scale large-kernel attention; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index measure.

To further illustrate the rationale and effectiveness of the proposed network design, we applied the four methods to the field seismic data for testing. Specifically, we selected a sample from the Netherlands F3 seismic survey as the low-resolution input for the network prediction. The prediction results of each network are shown in Figure 8.

The yellow dashed boxes in Figure 8 highlight each model's capability to recover stratigraphic structures. In Figure 8B-D, the recovered strata appear overly smooth, leading to a loss of detailed information. In contrast, Figures 8E and F demonstrate the recovery of complex stratigraphic structures, preserving richer details. The regions indicated by the yellow arrows within the red solid boxes in Figure 8 show that these areas are reconstructed as discontinuous strata (Figure 8B-D). However, as shown in Figure 8A, these positions should represent continuous stratigraphic structures, accurately depicted in Figure 8E and F. Compared with Figure 8E, the super-resolution result in Figure 8F demonstrates that the proposed approach achieves satisfactory outcomes even with fewer DFE stages.

To further compare the reconstruction performance of different networks, we used field data from the Kerry

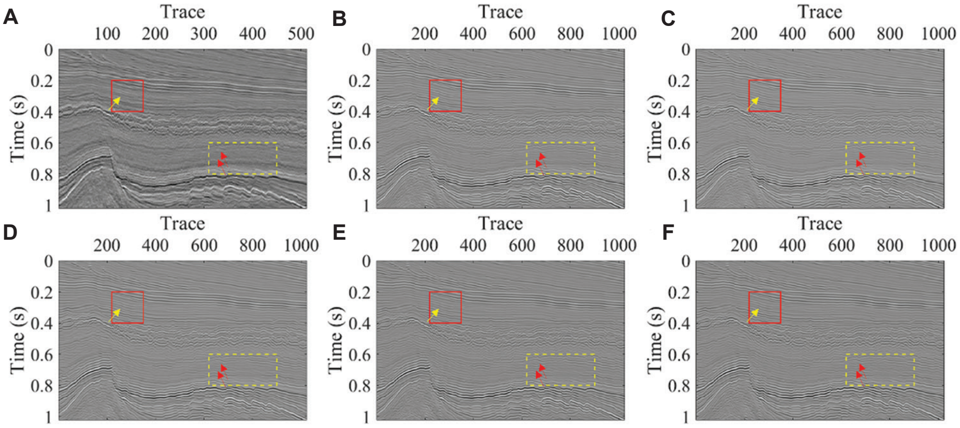


Figure 8. F3 field data prediction for networks with different structures. (A) Low-resolution field seismic image. (B) Super-resolution result by the grouped residual attention network. (C) Super-resolution result by $MLKA_2$. (D) Super-resolution result by $MLKA_3$. (E) Super-resolution result by $MLKA_5$. (F) Super-resolution result by $MLKA_4$.
Note: The subscript attached to "MLKA" denotes the number of deep feature extraction stages.
Abbreviation: MLKA: Multi-scale large-kernel attention.

block in New Zealand and extracted 512 traces for testing. Figure 9 displays the super-resolution results of these networks.

The yellow dashed boxes in Figure 9 highlight the stratigraphic recovery results of the different networks. The positions indicated by the red arrows reveal that Figure 9B-D fail to clearly delineate the detailed information of the stratigraphic structures, resulting in overly smooth reconstructions. In contrast, Figures 9E and F successfully recover each stratigraphic layer with greater clarity. The red solid boxes indicate the fault recovery regions, with the yellow arrows pointing to small fault structures. Figure 9B-D fail to recover these minor fault structures, instead reconstructing them as continuous stratigraphic layers with folded structures. Conversely, Figure 9E and F exhibit the highest accuracy in fault recovery.

Overall, the test and comparative analyses based on field data demonstrate that MLKA is crucial for improving the resolution of seismic images and reducing artifacts. Furthermore, the combined performance across different

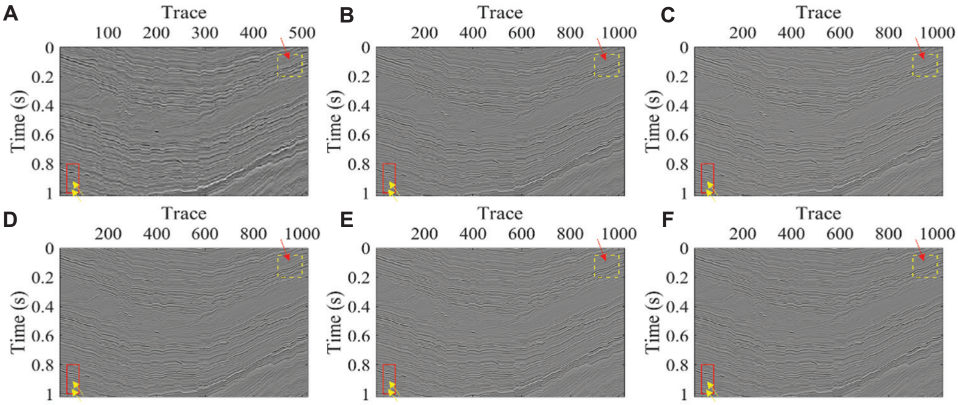


Figure 9. Kerry field data prediction for networks with different structures. (A) Low-resolution field seismic image. (B) Super-resolution result by the grouped residual attention network. (C) Super-resolution result by MLKA$_2$. (D) Super-resolution result by MLKA$_3$. (E) Super-resolution result by MLKA$_5$. (F) Super-resolution result by the proposed MLKA$_4$.
Note: The subscript attached to "MLKA" denotes the number of deep feature extraction stages.
Abbreviation: MLKA: Multi-scale large-kernel attention.

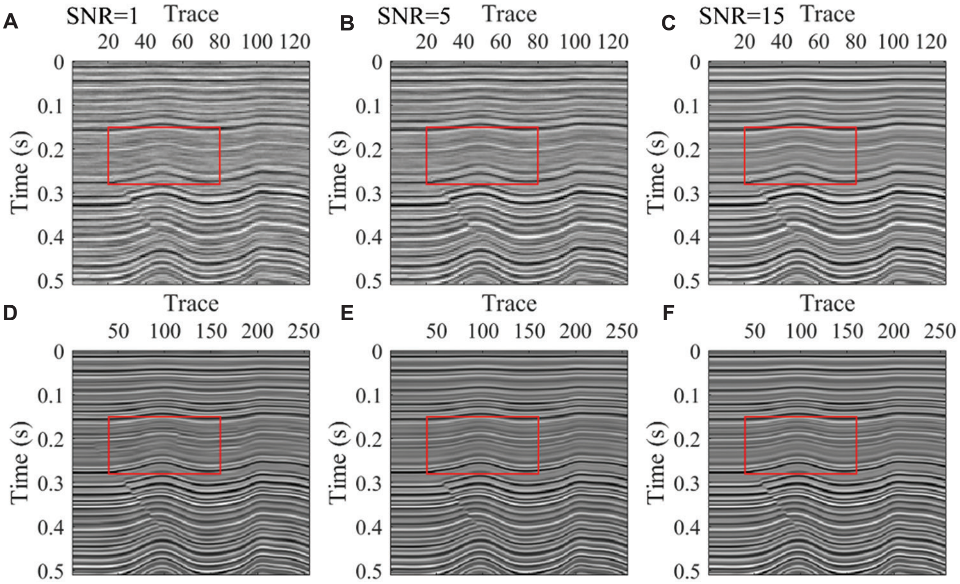


Figure 10. Comparison of denoising results for noisy synthetic seismic data at different signal-to-noise ratio (SNR) levels. (A) Noisy seismic data with SNR = 1 dB. (B) Noisy seismic data with SNR = 5 dB. (C) Noisy seismic data with SNR =15 dB. (D) Denoised result corresponding to panel A. (E) Denoised result corresponding to panel B. (F) Denoised result corresponding to panel C.

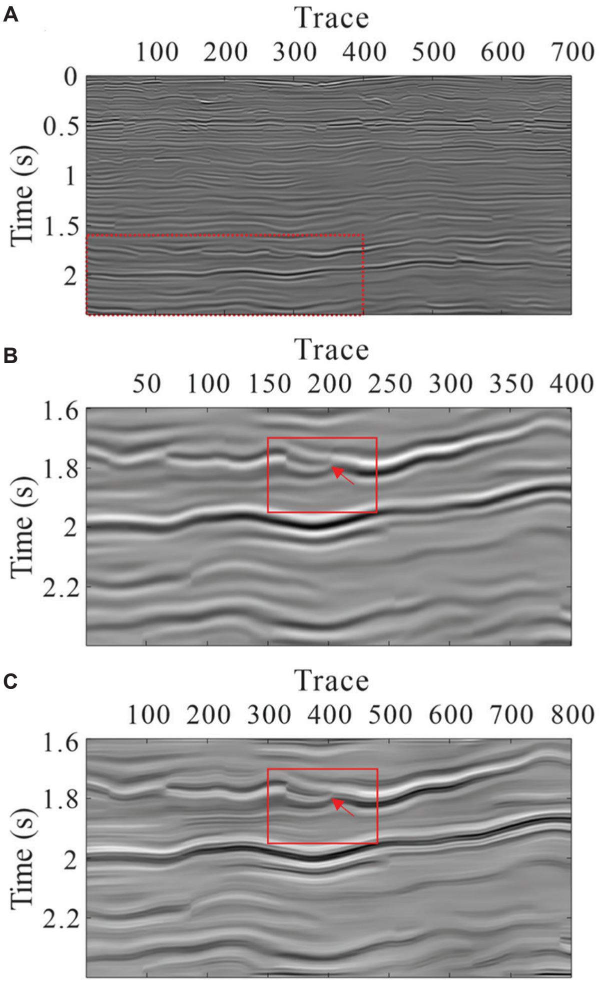


Figure 11. Test results of deep seismic data super-resolution. (A) Seismic data from the North Sea Volve Field in Norway. (B) Low-resolution deep seismic data. (C) Super-resolution result generated by the proposed approach.

DFE stage configurations suggests that a four-stage setup in the DFE component offers the best balance between accuracy and efficiency. This ablation study validates the interpretability and robustness of the proposed method.

4.2. Testing on noise robustness

We performed noise robustness assessments using synthetic seismic data engineered to exhibit distinct signal-to-noise ratio (SNR) characteristics. Specifically, we subjected high-resolution synthetic seismic data to a process of simple downsampling, followed by the introduction of colored noise at different intensities, as illustrated in Figure 10. The data represent SNR levels of 1 dB, 5 dB, and 15 dB, corresponding to strong, medium, and weak noise conditions, respectively.

On examination of the red solid-line boxes in Figure 10D-F, it is evident that the denoising performance at the strong noise level (1 dB) is less effective compared to the medium (5 dB) and weak (15 dB) noise levels. Notably, Figure 10D exhibits discontinuities in layer boundaries and the presence of artifacts, indicative of the challenges posed by high noise levels. Although the performance in low-SNR scenarios has not yet reached an optimal level, the majority of geological strata, including faults, have been successfully and clearly reconstructed. Figure 10E and F demonstrate a clearer restoration of the layer structure, highlighting the enhanced denoising capabilities of the proposed model under conditions of moderate and low noise, respectively. This comparative analysis underscores the adaptability and robustness of the proposed denoising technique in mitigating noise and preserving seismic data integrity across a spectrum of noise environments.

4.3. Testing on deep structures

In the preceding sections, we conducted super-resolution tests on field seismic data and denoising tests on synthetic seismic data. To showcase the adaptability of the proposed methodology, we applied it to the super-resolution of deep seismic data. In this segment, we selected deep seismic data from the North Sea Volve Field in Norway for the experiment, as depicted in Figure 11. The red dashed box in Figure 11A delineates the specific region of the deep seismic image that was extracted. This region is depicted in Figure 11B. The results illustrated in the figure indicate that the proposed model is effective in restoring layer boundaries. Within the red solid box, the minor fault indicated by the red arrow has also been clearly restored.

5. Conclusion

In this study, we propose GMLAN to achieve improved super-resolution for seismic images by integrating the benefits of group residual learning and large-kernel attention mechanisms. GMLAN comprises two primary components: feature extraction and image super-resolution reconstruction. Following the feature extraction phase, residual connections were employed to integrate the geological morphology and ground inclination characteristics of the seismic images. Subsequently, the output was fed into the IRM to produce super-resolution seismic images. A suite of comparative and ablation experiments demonstrated that the proposed network can significantly enhance and restore the high-frequency details of seismic images while preserving low-frequency information, indicating that the proposed approach is both highly effective and precise for seismic image super-resolution. Furthermore, the proposed method demonstrated superior noise robustness under different noise conditions. However, its efficacy in low SNR scenarios has yet to be optimized. In future work, the authors plan to address this issue through two approaches: (i) implementing a multi-scale framework for adaptive noise analysis and suppression tailored to seismic data across various SNR levels; and (ii) developing an adaptive SNR enhancement strategy that fine-tunes super-resolution reconstruction parameters based on the input seismic images' SNR. By applying these strategies, the authors aim to substantially boost the proposed model's performance in low SNR environments and further enhance the super-resolution of seismic images.

Acknowledgments

None.

Funding

This work was supported by BGP's Science and Technology project (01-04-02-2024).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Anxin Zhang, Zhenbo Guo
Formal analysis: Shiqi Dong, Zhiqi We
Investigation: Zhiqi Wei
Methodology: Shiqi Dong, Zhenbo Guo
Software: Anxin Zhang, Shiqi Dong
Validation: Anxin Zhang, Zhenbo Guo
Visualization: Zhenbo Guo, Zhiqi Wei
Writing–original draft: Anxin Zhang, Zhenbo Guo
Writing–review & editing: Shiqi Dong, Zhiqi Wei

Availability of data

The data generated and analyzed during this study are available from the corresponding author on reasonable request.

References

1. Soubaras R, Dowle R, Sablon R. Broadseis: Enhancing interpretation and inversion with broadband marine seismic. *CSEG Recorder*. 2012:37(7):40-46.

2. Rebert T, Sablon R, Vidal N, Charrier P, Soubaras R. Improving pre-salt imaging with variable-depth streamer data. In: *SEG Technical Program Expanded Abstracts 2012*. United States: Society of Exploration Geophysicists; 2012. p. 1-5.

   doi: 10.1190/segam2012-1067.1

3. Wang Y, Wang J, Wang X, Sun W, Zhang J. Broadband processing key technology research and application on slant streamer. *International Geophysical Conference, Beijing, China, 24-27 April 2018. Society of Exploration Geophysicists and Chinese Petroleum Society*. United States: Society of Exploration Geophysicists; 2018. p. 135-138.

   doi: 10.1190/IGC2018-034

4. Zhang YG, Wang Y, Yin JJ. Single point high density seismic data processing analysis and initial evaluation. *Shiyou Diqiu Wuli Kantan Oil Geophys. Prospect*. 2010;45:201-207.

   doi: 10.13810/j.cnki.issn.1000-7210.2010.02.008

5. Xiao F, Yang J, Liang B, *et al*. High-density 3D point receiver seismic acquisition and processing - a case study from the Sichuan Basin, China. *First Break*. 2014;32(1):81-90.

   doi: 10.3997/1365-2397.32.1.72598

6. Zhang H, Bao X, Zhao H, *et al*. High-precision deblending of 3-D simultaneous source data based on prior information constraint. *IEEE Geosci Remote Sens Lett.* 2025;22:1-5.

   doi: 10.1109/LGRS.2025.3526972

7. Shang XM, Diao R, Feng YP, Zhao CX. The application of spectral modeling method to high resolution processing of seismic data. *Geophys Geochem Explor*. 2014;38(1):75-80.

   doi: 10.11720/j.issn.1000-8918.2014.1.13

8. Wang D, Yuan S, Liu T, Li S, Wang S. Inversion-based non-stationary normal moveout correction along with prestack high-resolution processing. *J Appl Geophy*. 2021;191:104379.

   doi: 10.1016/j.jappgeo.2021.104379

9. Wu X, Ma J, Si X, *et al*. Sensing prior constraints in deep neural networks for solving exploration geophysical problems. *Proc Natl Acad Sci U S A*. 2023;120(23):e2219573120.

   doi: 10.1073/pnas.2219573120

10. Mousavi SM, Beroza GC, Mukerji T, Rasht-Behesht M. Applications of deep neural networks in exploration seismology: A technical survey. *Geophysics*. 2023;89(1):WA95-WA115.

    doi: 10.1190/geo2023-0063.1

11. Yuan S, Yu Y, Sang W, Xie R, Zhou C, Chen S. Seismic horizon picking using deep learning with multiple attributes. *IEEE Trans Geosci Remote Sens.* 2025;63:1-16.

    doi: 10.1109/TGRS.2025.3581462

12. Zeng D, Xu Q, Pan S, Song G, Min F. Seismic image super-resolution reconstruction through deep feature mining network. *Appl Intell*. 2023;53(19):21875-21890.

    doi: 10.1007/s10489-023-04660-y

13. Zhou R, Zhou C, Wang Y, Yao X, Hu G, Yu F. Deep learning with fault prior for 3-D seismic data super-resolution. *IEEE Trans Geosci Remote Sens.* 2023;61:1-16.

    doi: 10.1109/TGRS.2023.3262884

14. Li J, Wu X, Hu Z. Deep learning for simultaneous seismic image super-resolution and denoising. *IEEE Trans Geosci*

Remote Sens. 2022;60:1-11.

doi: 10.1109/TGRS.2021.3057857

15. Min F, Wang L, Pan S, Song G. D2UNet: Dual decoder U-net for seismic image super-resolution reconstruction. *IEEE Trans Geosci Remote Sens*. 2023;61:1-13.

    doi: 10.1109/TGRS.2023.3264459

16. Liang J, Cao J, Sun G, Zhang K, Van Gool L, Timofte R. SwinIR: Image Restoration using Swin Transformer. *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Montreal*, *BC*, *Canada*. 2021. p. 1833-1844.

    doi: 10.1109/ICCVW54120.2021.00210

17. Zhong T, Zheng K, Dong S, Tong X, Dong X. Enhancing the resolution of seismic images with a network combining CNN and transformer. *IEEE Geosci Remote Sens Lett*. 2025;22:1-5.

    doi: 10.1109/LGRS.2024.3495659

18. Zhong T, Yang F, Dong X, Dong S, Luo Y. SHBGAN: Hybrid bilateral attention GAN for seismic image super-resolution reconstruction. *IEEE Trans Geosci Remote Sens*. 2024;62:1-12.

    doi: 10.1109/TGRS.2024.3492142

19. Lin L, Zhong Z, Cai C, Li C, Zhang H. SeisGAN: Improving seismic image resolution and reducing random noise using a generative adversarial network. *Math Geosci*. 2024;56(4):723-749.

    doi: 10.1007/s11004-023-10103-8

20. Xiao Y, Li K, Dou Y, Li W, Yang Z, Zhu X. Diffusion models for multidimensional seismic noise attenuation and superresolution. *Geophysics*. 2024;89(5):V479-V492.

    doi: 10.1190/geo2023-0676.1

21. Yuan S, Xu Y, Xie R, Chen S, Yuan J. Multi-scale intelligent fusion and dynamic validation for high-resolution seismic data processing in drilling. *Pet Explor Dev*. 2025;52(3):680-691.

    doi: 10.1016/S1876-3804(25)60596-9

22. Zhou G, Zhi H, Gao E, *et al*. DeepU-Net: A parallel dual-branch model for deeply fusing multiscale features for road extraction from high-resolution remote sensing images. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2025;18:9448-9463.

    doi: 10.1109/JSTARS.2025.3555636

23. Zhou Y, Li Z, Guo CL, Bai S, Cheng MM, Hou Q. SRFormer: Permuted self-attention for single image super-resolution. In: *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. United States: IEEE; 2023. p. 12734-12745.

    doi: 10.1109/ICCV51070.2023.01174

24. Li Y, Deng Z, Cao Y, Liu L. GRFormer: Grouped residual self-attention for lightweight single image super-resolution. In: *Presented at: Proceedings of the 32nd ACM International Conference on Multimedia; 2024; Melbourne VIC, Australia*. United States: Cornell University.

    doi:10.1145/3664647.3681554

25. Yan Z, Zi-Xin W, Lin-qI C, Hong-Li D. Research on microseismic event localization based on convolutional neural network. *JSE*. 2024;33(6):1-32.

26. Xu Y, Yuan S, Zeng H, *et al*. Frequency-dependent multiscale network for seismic high-resolution processing. *Geophysics*. 2025;90(4):V297-V312.

    doi: 10.1190/geo2023-0682.1

27. Wang Y, Li Y, Wang G, Liu X. Multi-Scale Attention Network for single Image Super-Resolution. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; 2024. p. 5950-5960.

28. Wu X, Geng Z, Shi Y, Pham N, Fomel S, Caumon G. Building realistic structure models to train convolutional neural networks for seismic structural interpretation. *Geophysics*. 2019;85(4):WA27-WA39.

    doi: 10.1190/geo2019-0375.1

29. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE Trans Comput Imaging*. 2017;3(1):47-57.

    doi: 10.1109/TCI.2016.2644865

30. Lou Y, Zhang B, Wang R, Lin T, Cao D. Seismic fault attribute estimation using a local fault model. *Geophysics*. 2019;84(4):O73-O80.

    doi: 10.1190/geo2018-0678.1

31. Yan B, Wang T, Ji Y, Yuan S. Multidirectional coherence attribute for discontinuity characterization in seismic images. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

    doi: 10.1109/LGRS.2022.3151686

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Dual-branch dense network for seismic background noise elimination

Wei Wang[1], Haoliang Chen[2], Dekuan Chang[1], Xinyang Wang[3]*, Shujiang Wang[1], and Dong Li[1]

[1]Geophysics Institute, Research Institute of Petroleum Exploration and Development-Northwest, PetroChina, Lanzhou, Gansu, China

[2]Northeast Electric Power Design Institute CO., LTD. of China Power Engineering Consulting Group, Changchun, Jilin, China

[3]Department of Communication Engineering, College of Electric Engineering, Northeast Electric Power University, Jilin City, Jilin, China

(This article belongs to the *Special Issue: Advanced Artificial Intelligence Theories and Methods for Seismic Exploration)*

Abstract

Distributed acoustic sensing (DAS) has attracted much attention in seismic data acquisition because of its low cost, anti-electromagnetic interference, and high acquisition density. Unfortunately, the acquired DAS records are usually accompanied by various kinds of complex noise, affecting subsequent interpretation and inversion. Traditional methods have difficulties in effectively attenuating the intense background noise. In general, the denoising task of DAS data is challenging. Recently, convolutional neural networks (CNNs) exhibit a good ability in suppressing the noise in DAS records. However, traditional CNN-based frameworks always have a relatively simple network architecture, bringing negative impacts on the denoising capability. To solve this problem, we propose a dual-branch dense network (DBD-Net) in this paper. Specifically, DBD-Net introduces a novel combination of dual-branch modules and an attention mechanism: the dual-branch modules extract multi-scale coarse-to-fine features, while the attention mechanism highlights the most informative features. This joint design strengthens feature representation and signal recovery compared with conventional CNN structures such as denoising CNN (DnCNN) and U-Net. Moreover, an attention module is employed to enhance the effective features. To verify the denoising ability, we compare DBD-Net with other competing methods, including band-pass filter, DnCNN, and U-Net, in terms of denoising capability and processing accuracy. Experimental results verify that DBD-Net can improve the quality of DAS records with a signal-to-noise ratio increment of nearly 26 dB. Meanwhile, the intense DAS background noise is also perfectly suppressed and the weak signals are effectively restored, representing advantages over the competing methods.

Keywords: Background noise suppression; Distributed acoustic sensing; Convolutional neural network; Vertical seismic profile; Signal-to-noise ratio improvement

***Corresponding author:**
Xinyang Wang
(wxy23@neepu.edu.cn)

Citation: Wang W, Chen H, Chang D, Wang X, Wang S, Li D. Dual-branch dense network for seismic background noise elimination. *J Seismic Explor*.2025;34(5):53-65. doi: 10.36922/JSE025290038

Received: July 20, 2025

1st revised: September 18, 2025

2nd revised: October 23, 2025

Accepted: October 24, 2025

1. Introduction

Distributed acoustic sensing (DAS), by capturing the phase change of Rayleigh scattered light in optical fibers, is capable of accurately detecting subtle ground deformations induced by artificial seismic waves. Compared with traditional geophones, it exhibits advantages in acquisition properties and has been extensively used for seismic exploration, mainly combined with the vertical seismic profiles technique.[1-3] Unfortunately, effective signals acquired by the DAS system are relatively weak and are often susceptible to interference from strong background noise.[4] In addition, the DAS noise has complicated characteristics, and a few of them are not observed in the seismic records acquired by the geophones, such as fading noise, horizontal noise, and coupling noise.[5-8] It means that the traditional denoising methods might be unable to accurately attenuate this specific noise, leading to a degradation in denoising performance.[9,10] Therefore, effectively suppressing seismic noise and recovering effective signals has become a key aspect of seismic data processing.

Recently, researchers proposed many denoising methods to process seismic data, and some remarkable results have been achieved. In general, traditional denoising methods are classified into four categories: time-frequency analysis methods,[11,12] signal-decomposition-based methods,[13] multi-scale analysis methods,[14,15] and low-rank-based methods.[16,17] First, the time-frequency analysis methods can detect the effective signals by utilizing the differences in physical or spectral properties between signal and noise components. Typical methods include the band-pass filter (BPF),[18] median filter,[19] and Wiener filter.[20] Their denoising principles are relatively idealized. Thus, their performance might degrade when processing intense DAS background noise. Aiming to enhance the denoising ability, signal-decomposition-based methods, such as empirical mode decomposition,[21] variational mode decomposition,[22,23] and singular value decomposition,[24] were proposed. Generally speaking, these methods can decompose the signal into different modes and then accomplish the denoising task by retaining the signal-dominant modes. However, the signals and noise have overlaps, which limit the decomposition accuracy. Therefore, extracting the signal components from the mixed mode is often difficult and significantly affects the denoising performance. Meanwhile, multi-scale analysis methods are another type of commonly used denoising approach in the exploration industry. The application of these methods considers the fact that seismic signals and background noise have distinct characteristics in sparsity. In other words, we can use multi-scale analysis, such as shearlet,[14,25,26] seislet,[27] and curvelet transform,[28] to convert the temporal signals into a given sparse domain, and apply a 2D filter to maintain the signal components. Unfortunately, it is difficult to determine the critical parameters. Inappropriate thresholds or filters will inevitably cause signal leakage or residual noise. Furthermore, the low-rank-based approach is based on the assumption that the matrices of pure signals have low-rank properties. Rank is increased by the presence of DAS noise.[16,17] Therefore, the denoising task is equivalent to an optimization problem, which aims to obtain the minimal rank and output the attenuation results. Although low-rank-based methods can provide great denoising performance, the determination of optimal rank and huge computational cost still hinder the wide application of these methods. In summary, traditional denoising methods all have their shortcomings, and the attenuation accuracy cannot be ensured when faced with low signal-to-noise ratio (SNR) DAS data.

Nowadays, deep-learning-based methods are widely used in the field of seismic data processing.[29-32] Some successful applications have been reported in waveform inversion,[33,34] classification,[35,36] interpolation,[37] and migration.[38-40] As one of the hot topics in deep learning, the convolutional neural network (CNN) has been applied to DAS background noise attenuation. Dong *et al.*[9] proposed a method based on denoising CNNs (DnCNNs) to suppress the noise of seismic data, which optimizes the parameters and training data of DnCNNs so that they can adapt to the desert seismic data denoising environment and achieve adaptive blind denoising. The feasibility of CNNs in seismic data denoising was also verified through experiments.[41] Liu *et al.*[42] applied U-Net in seismic data noise suppression to achieve more effective seismic data processing through multi-scale extraction and channel expansion. Moreover, they verified the effectiveness of U-Net in seismic signal recovery and intense noise attenuation. These studies demonstrate the effectiveness of CNN in seismic data processing and show the superiority over traditional denoising methods at the same time. Its excellent performance is mainly benefited from the network structure and training data. The shortage of high-quality training data is certainly an important potential problem for CNN-based methods. It is well recognized that acquiring pure signals from field data is infeasible, posing a significant challenge in constructing signal training dataset. In addition, low-quality training data always impacts the signal amplitude preservation and weak signal recovery. Although carefully designed datasets can mitigate this issue, the fundamental limitation still lies in the network's ability to effectively extract and utilize the features of DAS data. This highlights the need for more advanced architectures.

To enhance the processing accuracy, a novel dual-branch dense network (DBD-Net) is proposed. Unlike previous CNN-based frameworks with relatively simple architectures that limit denoising capability, our method employs dual-branch modules to extract both coarse and fine features at multiple scales. In addition, a dense module is introduced to alleviate the gradient disappearance and improve feature fusion, while an attention mechanism is designed to enhance effective signals under complex noise conditions. With these improvements, the proposed network cannot only suppress intense DAS background noise but also preserve weak seismic signals more effectively. Therefore, compared with conventional CNN models such as DnCNN and U-Net, our method provides a more targeted solution for DAS denoising. The paper consists of four parts. Section 1 is the introduction, which introduces the significance of the study and the current status of the study. Section 2 is the method, which introduces the CNN network designed in this paper and the process of constructing a high-quality training set. Section 3 presents the results, which demonstrate the effectiveness of DBD-Net through the comparison experiments of synthetic and field records. Section 4 is the conclusion of this paper.

2. Methods

Figure 1 shows DBD-Net's network structure. In general, DBD-Net consists of the dual-branch module, the dense module, and the attention module. Specifically, DBD-Net has two network branches, a main scale, and a self-attention module. In the main scale, five convolutional layers are first applied to capture the initial features. Then, the dual-branch module, shown in Figure 2, is used to capture the multi-scale features. Meanwhile, the dense module transfers the shallow features, aiming to fuse them with deep features and improve the feature representation. In addition, a self-attention module is used to assist the effective feature extraction. Next, we will detail the specific descriptions of the network components.

2.1. DB module

Figure 2 presents the structure of the dual-branch module. The dual-branch module has double branches, designed to achieve multi-scale feature extraction. Specifically, the top branch focuses on capturing local fine features, while the bottom branch captures global coarse features. In the top branch, the feature map is first up-sampled by bilinear interpolation. Then, the dilated convolutional layers with different dilation rates (*i.e.*, 2, 3, 5), alternated with regular convolutional layers, are used to capture detailed local features in DAS data. Meanwhile, skip connections are used to transfer and fuse the features. Here, the application of dilated convolution rapidly enlarges the receptive field, and the utilization of different dilation rates maintains the diversity of captured features and avoids feature loss caused by the chessboard effect. In the bottom branch, the feature map is first down-sampled using max-pooling, which reduces its resolution to highlight coarse structures. Then, the same set of dilated convolutional layers, alternated with regular convolutional layers, is applied to capture global features. Skip connections are also employed to transfer and fuse features at this scale. Finally, the outputs of the top and bottom branches are added with the initial input of

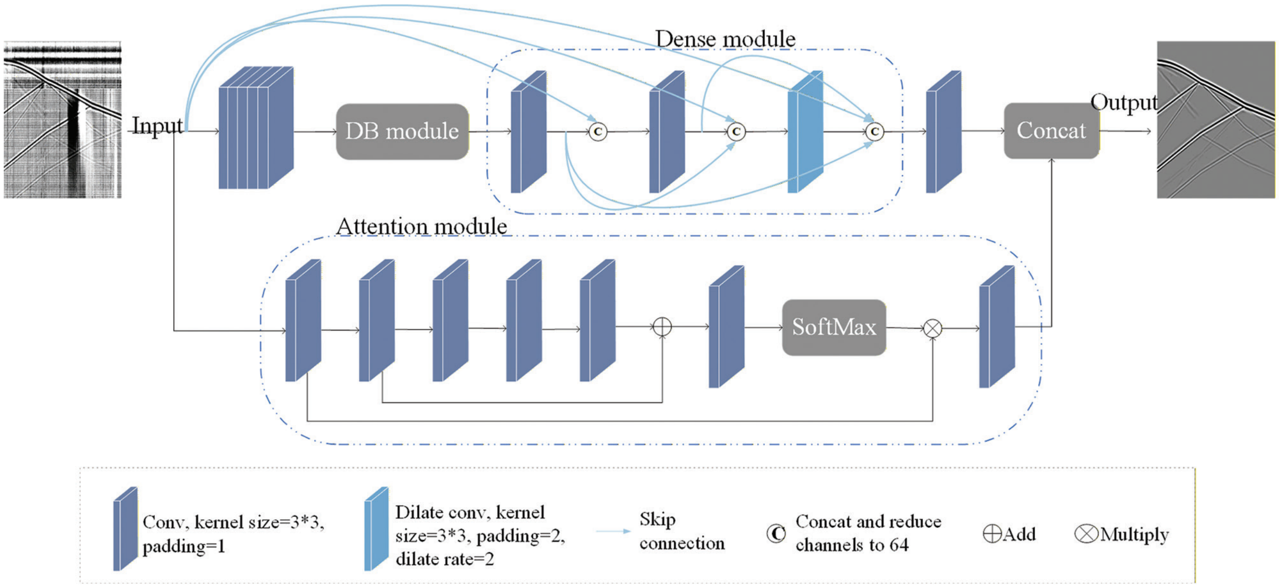


Figure 1. Structure of DBD-Net
Abbreviations: DB Module: Dual-branch module; DBD-Net: Dual-branch dense network.

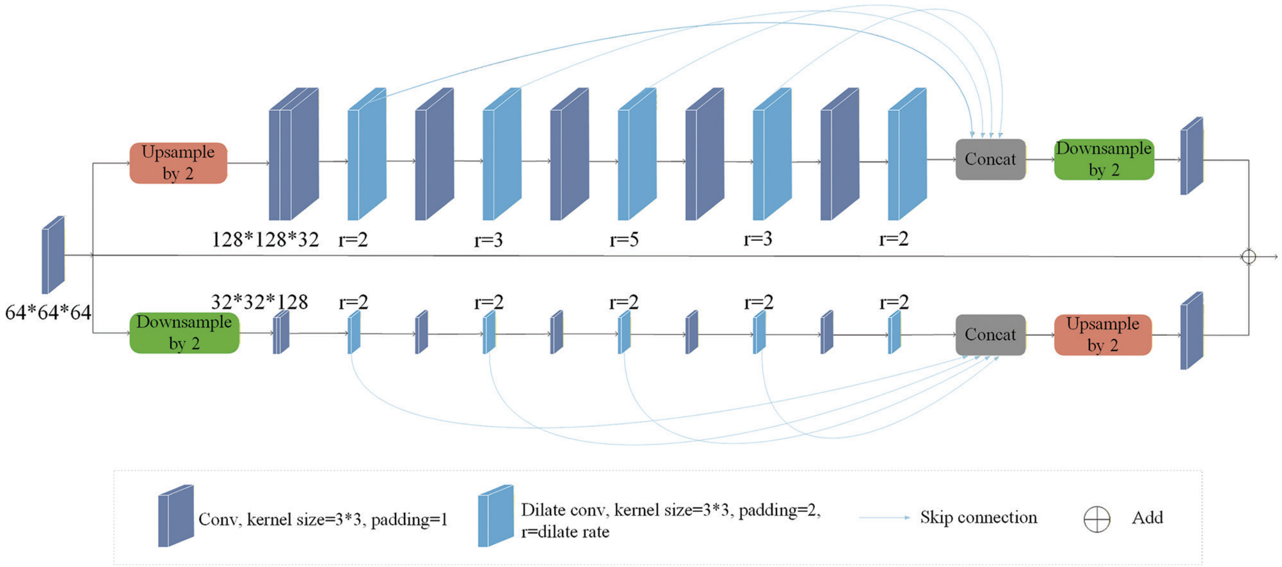


Figure 2. Structure of dual-branch module

the dual-branch module to enhance the effective features and reduce data redundancy.

2.2. Dense module

In general, with the depth of the network increasing, the learnable features are inundated with redundant information, which undoubtedly results in gradient disappearance and local optimum problems. The shallow features can reduce data redundancy and have good maintenance for gradients. In this study, we design a dense module, as depicted in Figure 1, to transfer and merge shallow features with deep features. The dense module has three convolutional layers, and all the outputs of convolutional layers are concatenated with the initial input data of the network through skip connections. It means that the feature utilization ability is enhanced and the gradient disappearance phenomenon is relieved, thereby improving the effective features.

2.3. Attention module

Effective signals in DAS data are commonly heavily affected by complex noise. To enhance the feature extraction capability, we design a spatial attention module. Specifically, the features are extracted using six convolutional layers, and then shallow feature fusion is implemented using skip connections. A softmax function is then used to activate the output of the sixth convolutional layer and distributions of the captured features. On this basis, the distributions are used as the weights to multiply with input features, and then, the desired features are enhanced because the effective features always have a larger probability value. Finally, the output of the attention module works as the guidance information to concatenate with the output of the main scale, thereby highlighting the primary features and improving the denoising accuracy, particularly for the weak up-going signals. In summary, the function of the attention module is to suppress noise-related features and emphasize effective seismic signals, significantly enhancing the recovery of weak events in complex DAS records.

2.4. Seismic denoising principle

In this study, we used DBD-Net to denoise noisy DAS records, and the detailed principle is shown below. The expression of the noisy DAS record is shown as follows:

$$y=x+n \tag{I}$$

where y denotes the noisy data, x and n represent the effective signals and background noise, respectively. Subsequently, y is used as the network input. The network output is an estimation of the effective signal, which is expressed as follows:

$$\hat{x}=F\left(y;\zeta\right) \tag{II}$$

where $\hat{x}$ denotes the estimation of the effective signal, F is the high-dimensional nonlinear mapping relationship constructed by the DBD-Net, with network parameters of $\zeta = \{\omega,k\}$. Here, ω is the weight, and k denotes bias. The network parameters are obtained randomly at the beginning, and the estimation error of the effective signal is quite large at this time. Then, we need to construct a loss function to evaluate error, which is defined as follows:

$$L_2(\zeta) = \frac{1}{N}\sum_{i=1}^{N}\left\| F(y_i;\zeta) - x_i \right\|^2 \tag{III}$$

Where y_i is the noisy record and x_i is the effective signal. We can iteratively optimize the parameters by using the backpropagation strategy. After training the network for several epochs, the error converges and becomes stable. In this study, we selected a set of parameters that minimize the loss function as the optimal set of parameters ζ_{out}. The final denoising results are given as follows:

$$\hat{x}_{out} \approx x = F(y;\zeta_{out}) \tag{IV}$$

2.5. Construction of the training set

Our training dataset contains a clean set of pure signal data and a set of field noise data. The pure signal dataset is simulated by simple seismic wavelets as in other previous studies, which have some differences from the properties of field DAS signals. It means that the performance of the trained models cannot be guaranteed and even recover some fake events in some conditions. To improve the denoising performance, we generated the clean signal training dataset using the forward modeling method. Specifically, it uses seismic wavelets to excite the artificial velocity model. Meanwhile, we utilized the finite difference method to simulate the propagation of the wavefield and then obtained synthetic data. To ensure the accuracy of learnable features, the previously acquired seismic profiles were considered references in the construction of velocity models. In this study, there were 20 velocity models. We obtained a total of 160 clean synthetic records. Table 1 presents the modeling parameters. The modeling equations of the Ricker Wavelet are described as:

$$L(t) = A \cdot \left[1 - 2\pi \cdot f(t - t_0)^2\right] \cdot e^{-\left[\pi \cdot f_0 (t - t_0)\right]^2} \tag{V}$$

where A is the amplitude, t_0 and f_0 denote the initial time and dominant frequency, respectively. As shown in Figure 3, we provide an example of a velocity model and synthetic recording. The inverted triangle at the top right identifies the source location, while the black line on the left symbolizes the receiving line, as shown in the figure. Subsequently, the synthetic records were divided into 64 × 64 patches. The signal dataset consisted of 20000 patches. For the field noise dataset, the passive source data acquired in Northeast China were employed. Similarly, the field noise data were divided into noise patches by the 64 × 64 sliding window. Here, we selected 30000 patches randomly to generate the test dataset. The field noise patches were multiplied by a weight ranging from 1 to 8 and then added with signal patches to form the noisy patches in the training process. We input signal patches and noise patches into DBD-Net, and through this training process, we successfully obtained denoising models. Figure 4 shows some samples of pure signal and field noise patches.

Table 1. Forward modeling parameters

Parameters	Value
Seismic wavelet	Ricker wavelet
Seismic wave frequency (Hz)	20–110
Wave velocity (m/s)	800–4700
Trace space (m)	1
Offset (m)	100–300
Sampling interval (s)	0.0004

2.6. Training details

In general, the high-performance computation technique affects CNN-based methods' performance. In this work, a graphics workstation was built with a configuration consisting of Intel (R) Core (TM) i5-9400F, NVIDIA GeForce RTX 2060 Super, and 16GB RAM. The specific parameters are listed in Table 2.

3. Results

3.1. Synthetic data results

Figure 5 shows the results concerning the synthetic DAS record, with Figure 5A presenting the corresponding velocity model, which has four layers, for the synthetic DAS record shown in Figure 5B. The fundamental frequency of the Ricker wavelet is 40 Hz.

Figure 5C shows the field seismic noise, which was combined with the synthetic DAS record to yield the noisy record, as shown in Figure 5D, which was used as the analyzed data. Here, we used BPF, DnCNN,[43] and U-Net,[44] for experimental comparison to process the noisy data with the proposed DBD-Net. Specifically, the pass band of BPF is 20–90 Hz, while DnCNN and U-Net use the same strategy and dataset as DBD-Net to train the denoising models.

The noisy data is shown in Figure 6A. BPF (Figure 6B) cannot effectively suppress the complex DAS noise, whose limited performance is demonstrated by the signal leakage and residual horizontal noise. Compared with BPF, CNN-based frameworks always provide better denoising results and the weak signals can also be restored. Among these methods, DnCNN and U-Net (Figure 6C and D) are outperformed by DBD-Net (Figure 6E). Notably, recovered events in DBD-Net are continuous and smooth, indicating its advantages in intensive DAS noise suppression. On this basis, two areas of interest are enlarged, as shown in Figure 7, for detailed comparisons. Area 1 is the local

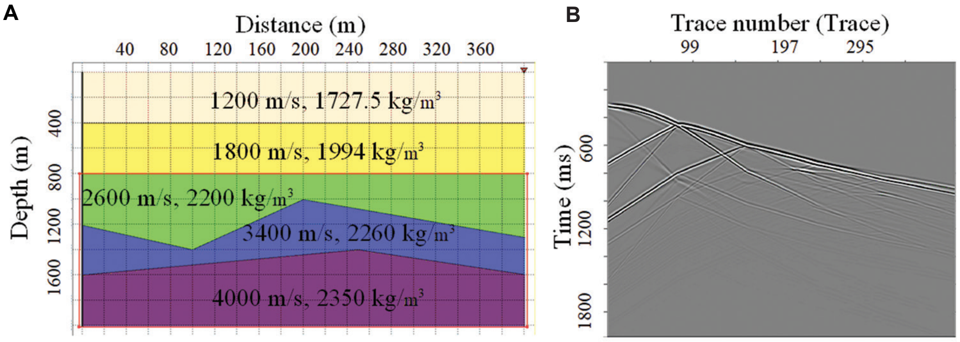


Figure 3. Clean synthetic record. (A) Forward model. (B) Synthetic record.

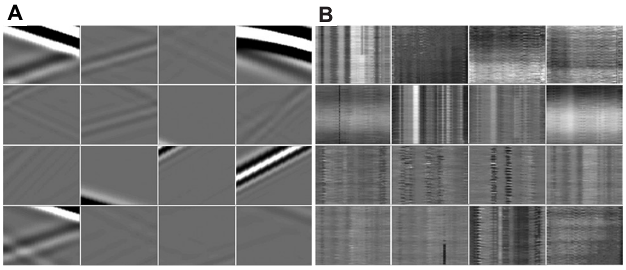


Figure 4. Samples of the training dataset. (A) Clean and noisy signal patches. (B) Noisy signal patches.

Table 2. Network architecture parameters

Hyper-parameter	DnCNN	U-Net	DBD-Net
Optimizer	ADAM	ADAM	ADAM
Patch size	64×64	64×64	64×64
Batch size	32	32	32
Epoch	60	60	60
Learning rate range	$[10^{-4},10^{-5},10^{-6}]$	$[10^{-4},10^{-5},10^{-6}]$	$[10^{-4},10^{-5},10^{-6}]$
Input channels	1	1	1
Layers	17	19	22
Convolution kernel size	3×3 × 64	3×3 × 64	3×3 × 64

Abbreviations: DBD-Net: Dual-branch dense network; DnCNN: Denoising convolutional neural network.

record (900–1200 traces) ranging from 1.5 to 1.8 s, which is contaminated by complex optical noise, while area 2 is the local record (240–540 traces) between 1.3 and 1.7 s, which is dominated by weak reflection events (Figure 7A). Similar results could be observed that BPF is affected by disordered information (Figure 7B), which also brings difficulty for the recognition of weak signals. Although DnCNN can effectively attenuate the background noise to some extent, the recovered events are corrupted and few weak signals have been recovered (Figure 7C). In addition, U-Net (Figure 7D) and DBD-Net represent close denoising ability; however, the recovered signals of DBD-Net are more complete and smoother (Figure 7E), indicating the impressive performance of DBD-Net in intensive DAS noise suppression.

Meanwhile, the F-K spectral analysis was conducted to compare the denoising method from the spectral perspective. As shown in Figure 8A, there is a conspicuous aliasing phenomenon between the clean signals and field noise. BPF (Figure 8B) shows limited effects when confronted with aliasing noise, and only noise components out of the pass band could be attenuated. The comparisons for the CNN-based frameworks can be achieved from two aspects: signal recovery performance and noise attention ability. On the one hand, there are obvious signal leakage components in the filtered noise results of DnCNN and U-Net (Figure 8C and 8D), demonstrating the adverse effect on signal amplitude preservation during the denoising process. On the other hand, the reconstructed results of DBD-Net (Figure 8E) have the most similar properties as the clean signals, indicating its capability in intensive DAS noise suppression. In summary, the comparisons in the F-K domain illuminate the superiority of DBD-Net.

In addition, a quantitative analysis was performed to further investigate the performance of each denoising method. In signal processing, SNR and root-mean-square-error (RMSE) are two key metrics for evaluating the denoising effect, and their specific expressions are as follows:

$$SNR = 10\lg\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} S_{ij}^{2}}{\sum_{i=1}^{n}\sum_{j=1}^{m}\left(N_{ij}-S_{ij}\right)^{2}}\right) \tag{VI}$$

$$RMSE = \sqrt{\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(N_{ij}-S_{ij}\right)^{2}} \tag{VII}$$

where S represents the clean record, and N represents the denoising result. In addition, i and j represent the

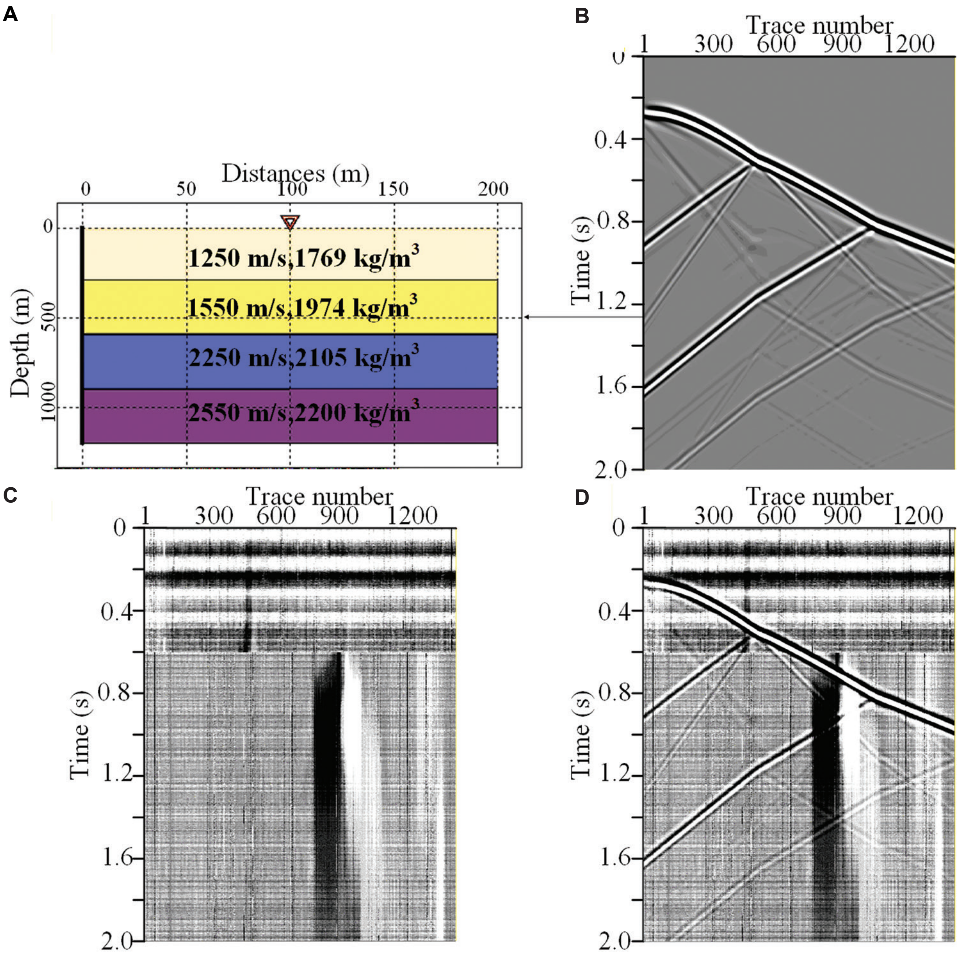


Figure 5. The synthetic DAS record for the test. (A) Forward model. (B) Synthetic clean seismic record. (C) Added noise. (D) Noisy DAS record with an SNR of 0 dB.

Abbreviations: DAS: Distributed acoustic sensing; SNR: Signal-to-noise ratio.

dimensions of the synthetic record. From the expressions, it can be seen that SNR is more inclined to measure the denoising effect, and a higher SNR means that the method is more capable of denoising. Meanwhile, RMSE is able to evaluate the effective signal maintenance ability, and a lower RMSE implies that the corresponding method is good at preserving the effective signal. To further investigate the denoising performance, we processed synthetic recordings with different SNRs. Table 3 presents the SNR and RMSE results for different methods. BPF has the lowest SNR increment among all the denoising methods, indicating its limited noise suppression ability. In contrast, the CNN-based methods have better performance, demonstrating significant SNR increment and small RMSE value. Under the same training environment, the training time of DnCNN, U-Net, and our proposed DBD-Net is about 1.26 h, 0.73 h, and 0.94 h, respectively. Although DBD-Net takes slightly longer to train than U-Net, the SNR improves by almost 3 dB. Compared with other methods, DBD-Net can increase the SNR by nearly 26 dB, reflecting its superiority in noise attenuation. Similar trends are also observed in RMSE results, indicating the effectiveness of DBD-Net in signal amplitude preservation.

3.2. Field data results

Figure 9 shows the results of the noise suppression. To better demonstrate the performance of DBD-Net in practical applications, the field DAS records in Figure 9A were processed using DBD-Net and other aforementioned methods. Specifically, the field DAS record, acquired in Northeast China, contains 1360 traces of seismic data. The corresponding sampling frequency is 2500 Hz. The reflection events in this field DAS record are severely corrupted by complex DAS interference, such as fading

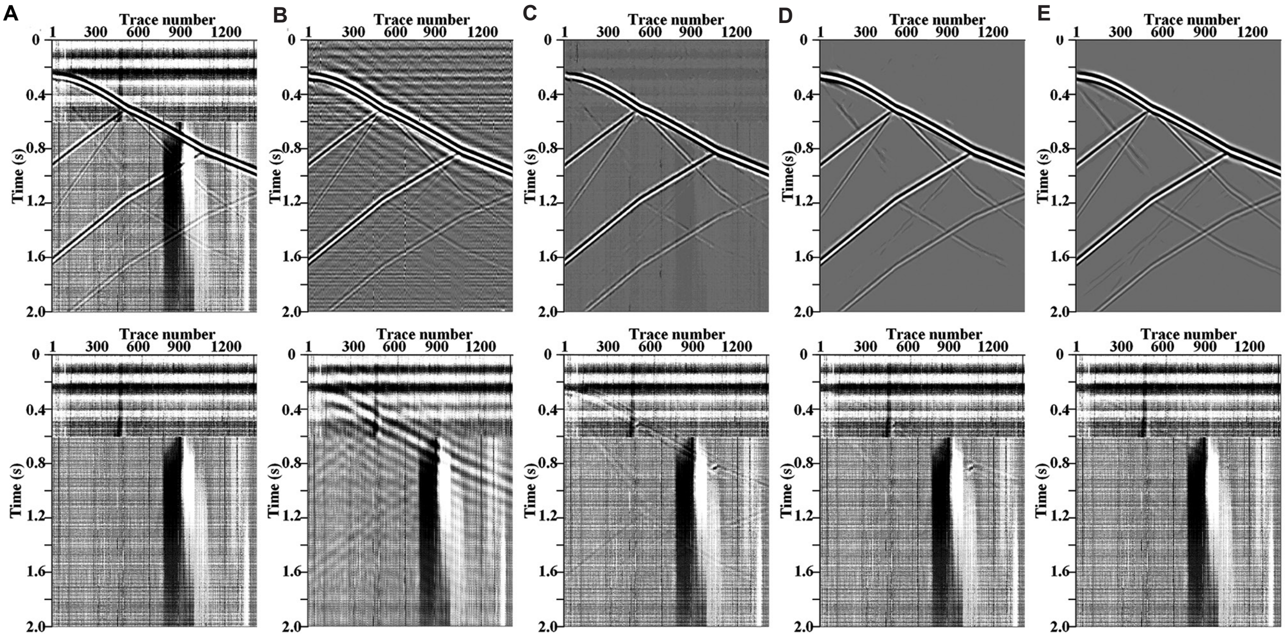


Figure 6. Synthetic record processing results. (A) Noisy record with an SNR of 0 dB and field data. (B) Results of BPF (10.24 dB). (C) Results of DnCNN (18.43 dB). (D) Results of U-Net (23.24 dB). (E) Results of DBD-Net (26.09 dB).

Abbreviations: BPF: Band-pass filter; DBD-Net: Dual-branch dense network; DnCNN: Denoising convolutional neural network; SNR: Signal-to-noise ratio.

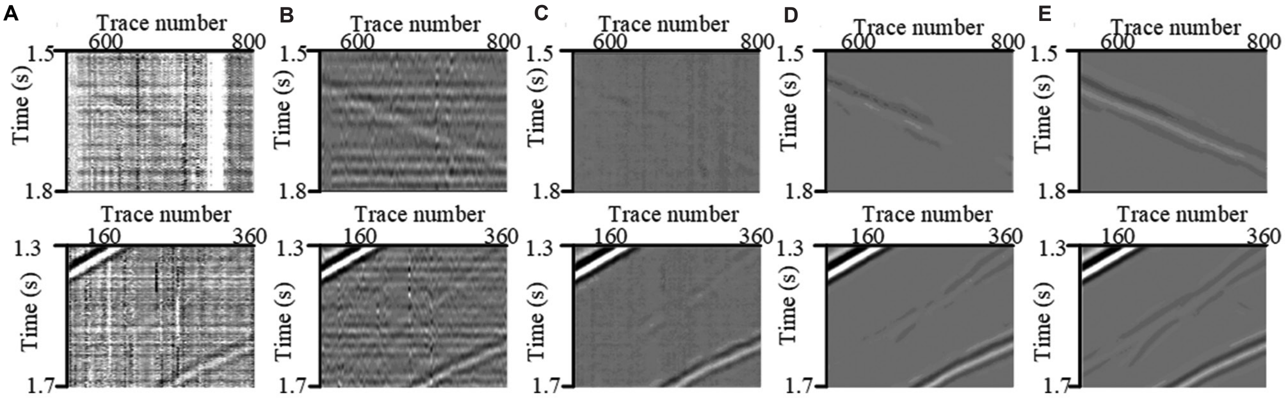


Figure 7. Enlarged figures showing denoising results corresponding to Figure 6. (A) Noisy record (top: Area 1, bottom: Area 2). (B–E) Results of BPF (B), DnCNN (C), U-Net (D), and DBD-Net (E) (top: Area 1, bottom: Area 2).

Abbreviations: BPF: Band-pass filter; DBD-Net: Dual-branch dense network; DnCNN: Denoising convolutional neural network.

noise and time-varying optical noise. It is known that BPF, as depicted in Figure 9B, cannot suppress the background noise in the pass band, resulting in residual horizontal noise and disordered recovered events. Moreover, as shown in Figure 9C, DnCNN has limited performance in DAS noise elimination, and the residual noise still influences the recognition of reflection signals, especially for weak up-going events. Although U-Net outperforms DnCNN (Figure 9D), the remained fading noise still affects the denoising results. In contrast, DBD-Net has a clear background with continuously recovered events (Figure 9E), indicating its great performance in denoising complex DAS data.

To further evaluate the denoising performance, we enlarged a local record (1000–1360 traces), ranging from 0.93 to 1.33 s as shown in Figure 10A. Similar to the results in Figure 9, BPF is incapable of suppressing the DAS background noise (Figure 10B). Meanwhile, as shown in

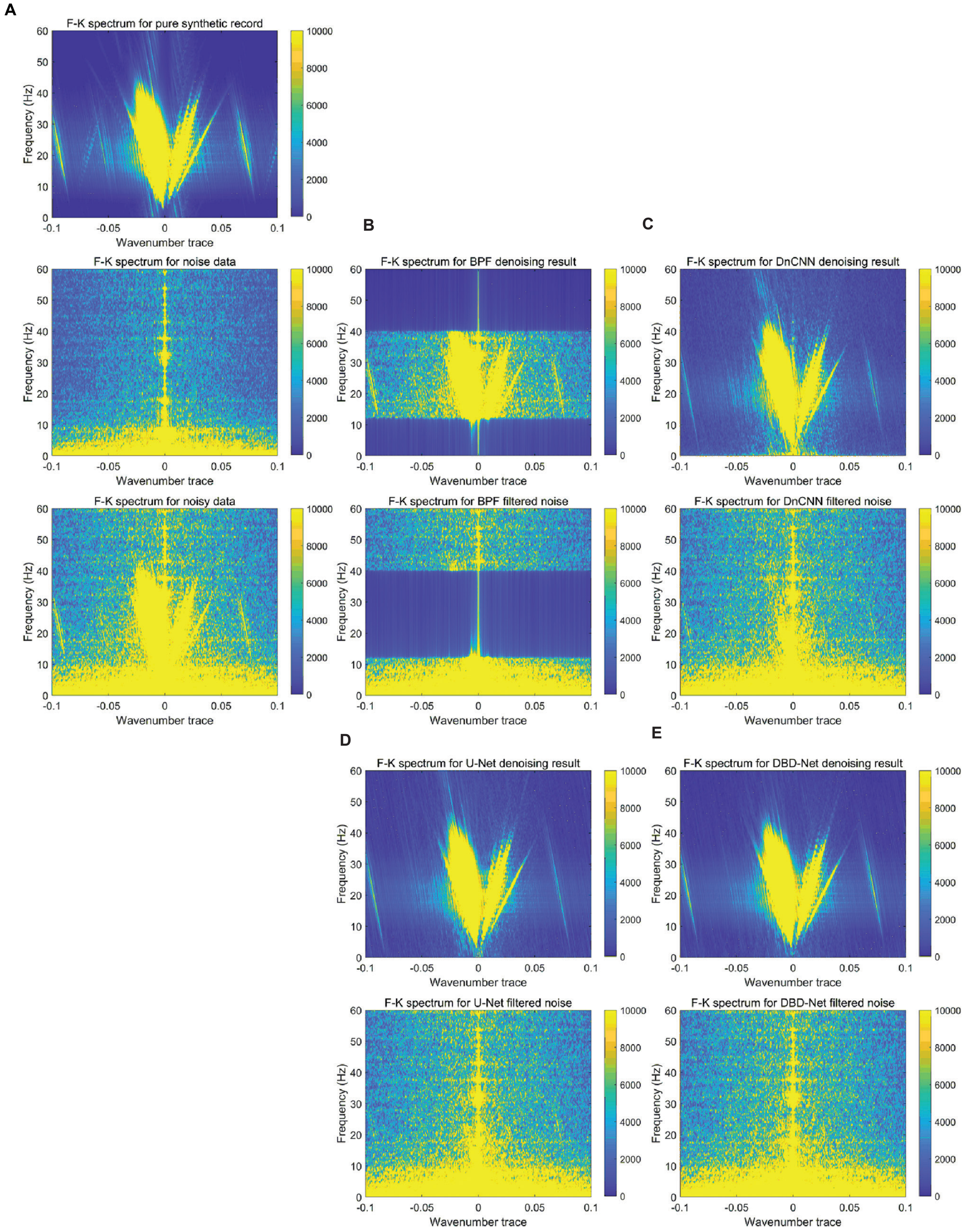


Figure 8. F-K spectral analysis. (A) F-K spectra of the clean record, noise data, and noisy data. (B–E) F-K spectra for the denoising results and filtered noise obtained by BPF (B), DnCNN (C), U-Net (D), and DBD-Net (E).

Abbreviations: BPF: Band-pass filter; DBD-Net: Dual-branch dense network; DnCNN: Denoising convolutional neural network.

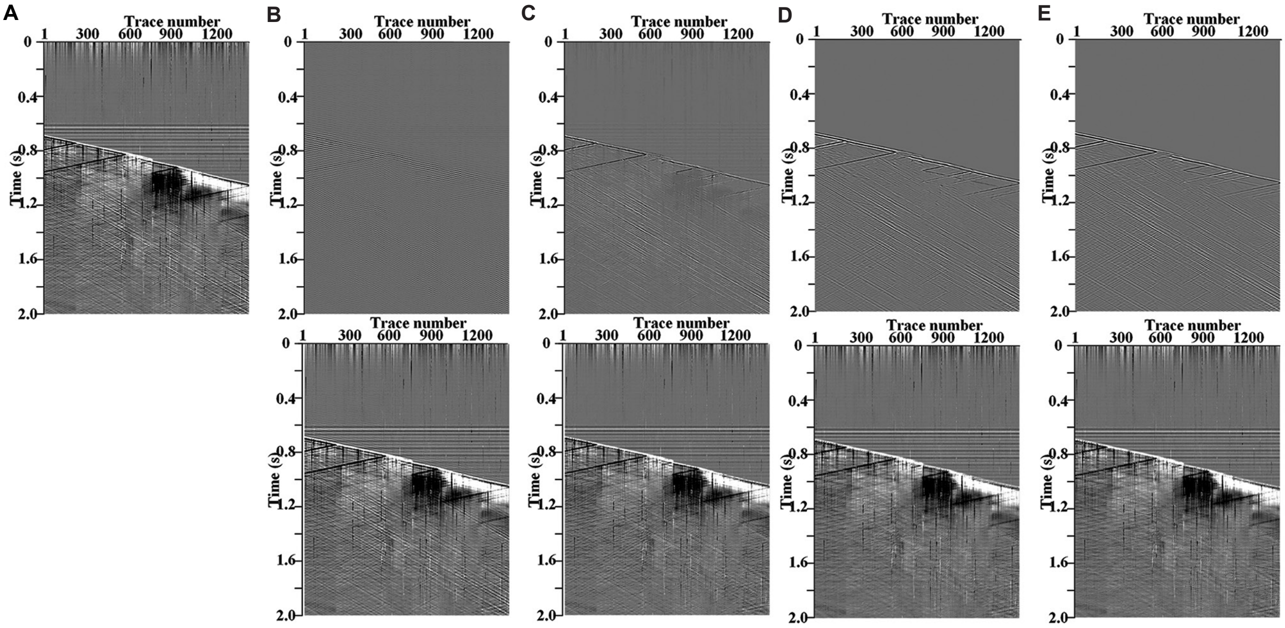


Figure 9. Field DAS records results. (A) Field DAS records. (b) Results of BPF. (C) Results of DnCNN. (D) Results of U-Net. (E) Results of DBD-Net.
Abbreviations: BPF: Band-pass filter; DAS: Distributed acoustic sensing; DBD-Net: Dual-branch dense network; DnCNN: Denoising convolutional neural network.

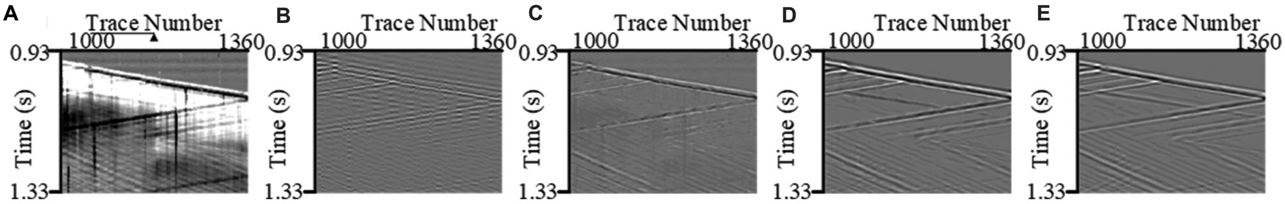


Figure 10. Enlargements of denoising results in Figure 9. (A) Field DAS records. (B) Results of BPF. (C) Results of DnCNN. (D) Results of U-Net. (E) Results of DBD-Net.
Abbreviations: BPF: Band-pass filter; DAS: Distributed acoustic sensing; DBD-Net: Dual-branch dense network; DnCNN: Denoising convolutional neural network.

Table 3. Comparison of SNR and RMSE among different algorithms

Synthetic record (dB)	BPF		DnCNN		U-Net		DBD-Net	
	SNR (dB)	RMSE	SNR (dB)	RMSE	SNR (dB)	RMSE	SNR (dB)	RMSE
0	10.24	0.2674	18.43	0.1041	23.24	0.0599	26.09	0.0431
−2	9.08	0.3056	17.27	0.1191	22.55	0.0648	25.46	0.0464
−4	7.71	0.3578	15.99	0.1378	21.70	0.0714	24.71	0.0505
−6	6.16	0.4277	14.58	0.1623	20.71	0.0807	23.81	0.0561
−8	4.47	0.5196	12.88	0.1973	19.60	0.0910	22.69	0.0638
−10	2.68	0.6387	10.94	0.2466	18.32	0.1055	21.36	0.0744

Abbreviations: BPF: Band-pass filter; DBD-Net: Dual-branch dense network; DnCNN: Denoising convolutional neural network; RMSE: Root-mean-square-error; SNR: Signal-to-noise ratio.

Figure 10C, DnCNN also performs below expectations, producing corrupted and discontinuous recovered signals. Among these methods, U-Net and DBD-Net, as displayed in Figure 10D and E, demonstrate similar denoising performance. However, DBD-Net is more competent in intense DAS noise elimination because of the clear background and smooth reconstructed events. Therefore, this analysis has proven the effectiveness of DBD-Net in denoising the complex DAS records, not only in noise attenuation but also in signal recovery.

4. Discussion

In this paper, we propose a CNN-based method named DBD-Net for seismic data denoising. Unlike traditional methods such as BPF, DBD-Net is capable of capturing complex seismic features and attenuating intense noise. Compared with representative CNN architectures such as DnCNN and U-Net, the main novelty of DBD-Net lies in the introduction of dual-branch modules for multi-scale feature extraction, together with a dense module for feature fusion and an attention mechanism for emphasizing effective signals. This design enables the network to more effectively restore weak seismic events while suppressing strong background noise.

Moreover, the noise dataset used in this study was collected from field DAS records and contains multiple types of interference, such as horizontal noise and fading noise. The experimental results show that DBD-Net maintains stable performance across these noise conditions, verifying its practical applicability. From a computational perspective, although training is relatively time-consuming due to the multi-branch design, the inference stage is efficient and suitable for practical denoising. In the future, lightweight strategies such as model pruning will be investigated to further reduce computational overhead and enhance deployment feasibility under different acquisition conditions.

5. Conclusion

In this study, we propose a novel multiscale denoising network that incorporates a multiscale scheme and an attention mechanism. Specifically, the proposed network uses a dual-branch module to extract coarse and detailed features in seismic data from multi-scale inputs. The attention module highlights and extends the primary features, thus improving the denoising ability. Meanwhile, for the training of this network, we have carefully designed a high-quality seismic training dataset. Synthetic clean records are constructed using the forward modeling method. The field seismic noise is superimposed on the dataset to guarantee the generalization of the training data. After processing the synthetic records and the field records, the results show that DBD-Net has a better performance in suppressing complex DAS background noise, coupled with an enhanced ability to accurately recover the effective signal, especially up-going signals with weak energy. Therefore, DBD-Net proves effective in suppressing strong DAS background noise and demonstrates notable application potential.

Acknowledgments

None.

Funding

This research was supported in part by the CNPC Science Research and Technology Development Project (2021DJ3505) and the Science and Technology Project of PetroChina Company Limited (2022KT1501).

Conflict of interest

Haoliang Chen is employed by Northeast Electric Power Design Institute CO., LTD. of the China Power Engineering Consulting Group. The remaining authors declare no conflict of interest.

Author contributions

Conceptualization: Wei Wang
Formal analysis: Wei Wang
Investigation: Wei Wang
Methodology: Wei Wang, Dekuan Chang
Software: Dong Li
Validation: Shujiang Wang
Writing–original draft: Haoliang Chen
Writing–review & editing: Xinyang Wang

Availability of data

The data analyzed for this study are included in the article. Further inquiries can be directed to the corresponding author.

References

1. Ashry I, Mao Y, Alias MS, *et al.* Normalized differential method for improving the signal-to-noise ratio of a distributed acoustic sensor. *Appl Opt*. 2019;58(18): 4933-4938.

   doi: 10.1364/AO.58.004933

2. Ashry I, Mao Y, Wang B, *et al.* A review of distributed fiber-optic sensing in the oil and GAS industry. *J Lightwave Technol*. 2022;40(5):1407-1431.

   doi: 10.1109/JLT.2021.3135653

3. Zhong T, Cheng M, Dong X, Li Y, Wu N. Seismic random noise suppression by using deep residual U-Net. *J Petrol Sci Eng*. 2022;209:109901.

   doi: 10.1016/j.petrol.2021.109901

4. Zhong T, Wang W, Lu S, Dong X, Yang B. RMCHN: A residual modular cascaded heterogeneous network for noise suppression in DAS-VSP records. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

   doi: 10.1109/LGRS.2022.3229556

5. Cooper HW, Cook RE. Seismic data gathering. *Proc IEEE*. 1984;72(10):1266-1275.

6. Zhang Z, Alajami M, Alkhalifah T. Wave-equation dispersion

spectrum inversion for near-surface characterization using fiber-optics acquisition. *Geophys J Int*. 2020;222(2):907-918.

doi: 10.1093/gji/ggaa211

7. Spikes KT, Tisato N, Hess TE, Holt JW. Comparison of geophone and surface-deployed distributed acoustic sensing seismic data. *Geophysics*. 2019;84(2):A25-A29.

   doi: 10.1190/geo2018-0528.1

8. Zhong T, Cheng M, Dong X, Wu N. Seismic random noise attenuation by applying multi-scale denoising convolutional neural network. *IEEE Trans Geosci Remote Sens*. 2022;60:590501.

   doi: 10.1109/TGRS.2021.3095922

9. Dong X, Li Y. Denoising the optical fiber seismic data by using convolutional adversarial network based on loss balance. *IEEE Trans Geosci Remote Sens*. 2021;59(12): 10544-10554.

   doi: 10.1109/TGRS.2020.3036065

10. Zhong T, Cheng M, Lu S, Dong X, Li Y. RCEN: A deep-learning-based background noise suppression method for DAS-VSP records. *IEEE Geosci Remote Sens Lett*. 2022;19:3004905.

    doi: 10.1109/LGRS.2021.3127637

11. Stockwell R, Manisinba L, Lowe R. Localization of the complex spectrum: The S transform. *IEEE Trans Signal Process*. 1996;44(4):998-1001.

    doi: 10.1109/78.492555

12. Liu N, Gao J, Zhang B, Wang Q, Jiang X. Self-adaptive generalized S-transform and its application in seismic time-frequency analysis. *IEEE Trans Geosci Remote Sens*. 2019;57(10):7849-7859.

    doi: 10.1109/TGRS.2019.2916792

13. Bekara M, van der Baan M. Random and coherent noise attenuation by empirical mode decomposition. *Geophysics*. 2009;74(5):V89-V98.

14. Guo K, Labate D. Optimally sparse multidimensional representation using shearlets. *SIAM J Math Anal*. 2007;39(1):298-318.

    doi: 10.1137/060649781

15. Houska R. The nonexistence of Shearlet scaling functions. *Appl Computat Harmon Anal*. 2012;32(1):28-44.

    doi: 10.1016/j.acha.2011.03.001

16. Anvari R, Siahsar M, Gholtashi S, Kahoo A. Seismic random noise attenuation using synchrosqueezed wavelet transform and low-rank signal matrix approximation. *IEEE Trans Geosci Remote Sens*. 2017;55(11):6574-6581.

    doi: 10.1109/TGRS.2017.2730228

17. Anvari R, Mohammadi M, Kahoo AR, Khan NA, Abdullah AI. Random noise attenuation of 2-D seismic data based on sparse low-rank estimation of the seismic signal. *Comput Geosci*. 2020;135:104376-104387.

    doi: 10.1016/j.cageo.2019.104376

18. Binder G, Titov A, Liu Y, *et al*. Modeling the seismic response of individual hydraulic fracturing stages observed in a time-lapse distributed acoustic sensing vertical seismic profiling survey. *Geophysics*. 2020;85(4):T225-T235.

    doi: 10.1190/GEO2019-0819.1

19. Wang Y, Liu X, Gao F, Rao Y. Robust vector median filtering with a structure-adaptive implementation. *Geophysics*. 2020;85(5):V407-V414.

    doi: 10.1190/GEO2020-0012.1

20. Mendel JM. White-noise estimators for seismic data processing in oil exploration. *IEEE Trans Automat Control*. 2003;AC-22(5):694-706.

    doi: 10.1109/TAC.1977.1101597

21. Han JJ, van der Baan M. Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding. *Geophysics*. 2015;80(6):69-80.

    doi: 10.1190/GEO2014-0423.1

22. Liu N, Li F, Wang D, Gao J, Xu Z. Ground-roll separation and attenuation using curvelet-based multichannel variational mode decomposition. *IEEE Trans Geosci Remote Sens*. 2022;60:5901214.

    doi: 10.1109/TGRS.2021.3054749

23. Yu S, Ma J. Complex variational mode decomposition for slope-preserving denoising. *IEEE Trans Geosci Remote Sens*. 2017;56(1):586-597.

    doi: 10.1109/TGRS.2017.2751642

24. Bekara M, van der Baan M. Local singular value decomposition for signal enhancement of seismic data. *Geophysics*. 2007;72(2):V59-V65.

25. Liu P, Li R, Yue YH, Liao SJ, Qian F. Robust prestack seismic facies analysis using shearlet transform-based deep learning. *J Geophys Eng*. 2022;19(3):521-533.

    doi: 10.1093/jge/gxac015

26. Liu C, Wang D, Sun J, Wang T. Crossline-direction reconstruction of multi-component seismic data with shearlet sparsity constraint. *J Geophys Eng*. 2018;15(5): 1929-1942.

    doi: 10.1088/1742-2140/aac097

27. Liu Y, Fomel S, Liu C. Signal and noise separation in prestack seismic data using velocity-dependent seislet transform. *Geophysics*. 2015;80(6):WD117-WD128.

    doi: 10.1190/GEO2014-0234.1

28. Neelamani R, Baumstein AI, Gillard DG, Hadidi MT, Soroka WL. Coherent and random noise attenuation using the curvelet transform. *Lead Edge*. 2008;27(2):240-248.

29. Li Y, Wang H, Dong X. The denoising of desert seismic data based on cycle-GAN with unpaired data training. *IEEE Geosci Remote Sens Lett*. 2021;18(11):2016-2020.

    doi: 10.1109/LGRS.2020.3011130

30. Dong X, Lin J, Lu S, Wang H, Li Y. Multiscale spatial attention network for seismic data denoising. *IEEE Trans Geosci Remote Sens*. 2022;60:21779964.

    doi: 10.1109/TGRS.2022.3178212

31. Dong X, Lin J, Lu S, Huang X, Wang H, Li Y. Seismic shot gather denoising by using a supervised deep learning method with weak dependence on real noise data: A solution to the lack of real noise data. *Surv Geophys*. 2022;43(5):1363-1394.

    doi: 10.1007/s10712-022-09702-7

32. Zhong T, Li F, Zhang R, Dong X, Lu S. Multi-scale residual pyramid network for seismic background noise attenuation. *IEEE Trans Geosci Remote Sens*. 2022;60:5922014.

    doi: 10.1109/TGRS.2022.3217887

33. Lin Y, Theiler J, Wohlberg B. Physics-guided data-driven seismic inversion: Recent progress and future opportunities in full-waveform inversion. *IEEE Signal Process Mag*. 2023;40(1):115-133.

    doi: 10.1109/MSP.2022.3217658

34. Chen G, Yang W, Wang H, Zhou H, Huang X. Elastic full waveform inversion based on full-band seismic data reconstructed by dual deconvolution. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

    doi: 10.1109/LGRS.2022.3178915

35. Li J, Ye M, Stankovic L, Stankovic V, Pytharouli S. Domain knowledge informed multitask learning for landslide-induced seismic classification. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

    doi: 10.1109/LGRS.2023.3279068

36. Noh K, Kim D, Byun J. Explainable deep learning for supervised seismic facies classification using intrinsic method. *IEEE Trans Geosci Remote Sens*. 2023;61:1-11.

    doi: 10.1109/TGRS.2023.3236500

37. Li X, Wu B, Zhu X, Yang H. Consecutively missing seismic data interpolation based on coordinate attention Unet. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

    doi: 10.1109/LGRS.2021.3128511

38. Lu S. Migration using sea surface-related multiples: Challenges and opportunities. *Geophysics*. 2021;86(5):WC11-WC19.

    doi: 10.1190/GEO2020-0862.1

39. Lu S, Wu H, Dong X, *et al*. Building adjoint operators for least-squares migration using the acoustic wave equation. *Geophysics*. 2023;88(2):S71-S85.

    doi: 10.1190/GEO2022-0279.1

40. Wang H, Lin J, Dong X, Li Y, Yang B. Seismic velocity inversion transformer. *Geophysics*. 2023;88(4):R513-R533.

    doi: 10.1190/GEO2022-0283.1

41. Dong X, Cheng M, Wang H, Li G, Lin J, Lu S. A potential solution to insufficient target-domain noise data: Transfer learning and noise modeling. *IEEE Trans Geosci Remote Sens*. 2023;61:5915115.

    doi: 10.1109/TGRS.2023.3300697

42. Liu N, Wang J, Gao J, Chang S, Lou Y. Similarity-informed self-learning and its application on seismic image denoising. *IEEE Trans Geosci Remote Sens*. 2022;60:1-13.

    doi: 10.1109/TGRS.2022.3210217

43. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans Image Process*. 2017;26(7):3142-3155.

    doi: 10.1109/TIP.2017.2662206

44. Olaf R, Philipp F, Thomas B. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer International Publishing; 2015. p234-241.

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Transposed arrangement strategy-based deep learning for seismic data crossline interpolation

Jiyun Yu[1†], Yonghwan Joo[2†], and Daeung Yoon[1*]

[1]Department of Energy and Resources Engineering, Chonnam National University, Gwangju, Republic of Korea

[2]Resources Exploration and Development Research Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), Daejeon, Republic of Korea

Abstract

Towed-streamer marine seismic acquisition systems generally have a dense receiver spacing in the inline receiver direction within common-shot gathers (along-streamer), while the streamer spacing is relatively sparse in the crossline receiver direction within common-shot gathers (cross-streamer). This disparity can lead to spatial aliasing issues in the crossline receiver direction within common-shot gathers and result in resolution degradation during the processing of 3D seismic data. To address this issue and enhance resolution, data interpolation in the crossline receiver direction within common-shot gathers is essential. Various supervised learning-based interpolation methods have been developed to this end. However, the absence of true data in the crossline receiver direction within common-shot gathers poses challenges for training supervised learning models with actual field data. To overcome this, we have developed a novel approach called the "transposed arrangement strategy" for a deep learning-based reconstruction model for crossline interpolation. This method involves training the model with 3D input and labels patched from existing field data, and then applying the trained model with transposed 3D input to reconstruct data in the crossline receiver direction within common-shot gathers. During this process, the 3D U-Net and U-Net+ models were utilized, demonstrating their superiority through comparisons with traditional interpolation methods.

Keywords: Data processing; Signal processing; Deep learning; Crossline interpolation; Transposed arrangement strategy; 3D U-Net+

1. Introduction

Seismic survey data can be subject to loss due to environmental, economic constraints, or mechanical defects. Notably, the towed-streamer marine seismic systems are typically configured with dense inline and wide crossline spacing to maximize exploration coverage.[1] This setup can lead to spatial aliasing issues in the crossline receiver direction within common-shot gathers and resolution degradation during the processing of 3D seismic survey data. Consequently, data interpolation in the crossline receiver direction within common-shot gathers is necessary to address these problems and improve data resolution.

Before the introduction of deep learning, a wide range of conventional approaches were developed for seismic data interpolation. Traditional methods include bicubic

---

†These authors contributed equally to this work.

***Corresponding author:**
Daeung Yoon
(duyoon@jnu.ac.kr)

Citation: Yu J, Joo Y, Yoon D. Transposed arrangement strategy-based deep learning for seismic data crossline interpolation. *J Seismic Explor*. 2025;34(5):66-80. doi: 10.36922/JSE025330056

Received: August 13, 2025

1st revised: September 22, 2025

2nd revised: October 24, 2025

Accepted: November 3, 2025

interpolation,[2] wave-equation–based methods,[3,4] and f–x interpolation methods.[5] Prediction error filtering[6] has also been widely applied for reconstructing missing traces by exploiting predictable structures in seismic data. In addition, projection onto convex sets (POCS)[7] methods and their extensions, such as curvelet-domain POCS,[8,9] have been shown to effectively exploit sparsity in transform domains. Rank reduction methods[10-12] and sparse transform-based methods, including Fourier, curvelet, and seislet transforms,[13-15] have also been extensively studied. These conventional methods perform well for linear or near-linear seismic events and can mitigate aliasing, but they often involve high computational cost, require careful parameter tuning (*e.g.*, window size, rank selection), and may struggle with complex or curved structures. With the increasing size and complexity of modern 3D seismic surveys, these limitations have motivated the exploration of more efficient and adaptive approaches, including deep learning–based methods.

Recent advancements in deep learning technology have been actively incorporated into seismic data interpolation research to provide high-accuracy interpolation results and more efficient data processing.[16-20] Research in deep learning-based seismic data interpolation can be categorized into techniques for 2D and 3D seismic data based on the dimensionality of the data.

2D seismic data interpolation techniques are primarily based on convolutional neural networks (CNNs), generative adversarial networks (GANs) and diffusion models. Yu and Wu[21] proposed a CNN that utilizes a hybrid loss function combining structural similarity index (SSIM) and L1 norm, along with an attention mechanism that explicitly leverages global information. Li *et al.*[22] proposed a method of integrating a coordinate attention block into U-Net for 2D successive missing traces interpolation. Park *et al.*[20] proposed a strategy using the Coarse-Refine U-Net (CFunet), which consists of two U-Nets and an upsampling process between them, along with the Fourier loss. Lou *et al.*[23] proposed a wavelet-based convolutional block attention deep learning (W-CBADL) network for the reconstruction of irregularly sampled seismic data. Dodda *et al.*[24] proposed the use of an attention-based wavelet convolutional neural network (AWUN) for simultaneous noise reduction and reconstruction of incomplete seismic data. Tian *et al.*[25] proposed feature restoration-based U-Net (FR–U-Net), a feature restoration-based interpolation method built upon the U-Net architecture to reconstruct consecutively missing seismic traces. Kaur *et al.*[26] proposed a model using the cycle GAN structure that comprises two generators and one discriminator. Chang *et al.*[27] proposed a method for interpolating seismic data using the conditional generative adversarial network in the time and frequency domains (TF-CGAN). Deng *et al.*[28] proposed conditional constraint diffusion model, a diffusion model with conditional constraints. However, these proposed methods are limited to interpolating 2D seismic data and have not been applied to multidimensional seismic data.

Consequently, 3D seismic data interpolation techniques have been developed, which predominantly involve models based on CNNs, GANs, diffusion models and additionally, autoencoder-based models. Kong *et al.*[29] proposed a multi-resolution U-Net model that utilizes the correlations in 3D data. Jin *et al.*[30] proposed a method to reduce the operational cost of U-Net by introducing depthwise separable convolution instead of standard convolution. Chang *et al.*[31] proposed a dual-domain conditional generative adversarial network that uses seismic data sets and discrete Fourier-transformed data sets in the frequency domain as input vectors. Dou *et al.*[32] proposed a multi-dimensional adversarial GAN that uses three discriminators. Yu and Yoon[33] applied the conditional Wasserstein GAN (cWGAN) model to 3D seismic data interpolation. Ding *et al.*[34] proposed Self-Attention Generative Adversarial Network, a deep learning-based model that integrates the self-attention mechanism with GAN. Wang *et al.*[35] proposed SeisFusion, a diffusion model combined with conditional constraints. Qian *et al.*[36] introduced the deep tensor autoencoder model, which is capable of learning data-driven, non-linear, and high-dimensional mappings.

Despite these technological advancements, these models often fall short of generating new data to reduce the spacing between traces compared to the original dataset. Typically, they demonstrate strong performance when interpolating from decimated data back to its full original format. This is because a supervised learning-based interpolation model requires a label, which is selected from the original dataset. For instance, if the original data has a spacing of 20 m, a model might be effectively trained to fill in missing traces corresponding to the label, thus decreasing the spacing from 40 m to 20 m by reconstructing the decimated data. However, this training does not guarantee a reduction in spacing from 20 m to 10 m, which is the ultimate goal of seismic data interpolation.

This limitation, however, can be mitigated in cross-streamer wavefield reconstruction due to the denser inline sensor spacing compared to the crossline receiver direction within common-shot gathers. In this scenario, an interpolation model can be trained along the inline receiver direction within common shot gathers and subsequently applied to reconstruct the coarser data in the crossline receiver direction within common-shot gathers. Larsen Greiner *et al.*[37] introduced a cross-streamer wavefield reconstruction model using a wavelet domain approach, training the model along the inline receiver direction within common-shot gathers to reconstruct coarse crossline receiver direction within common-shot gathers. Yeeh *et al.*[38] proposed a trace-to-trace approach for crossline interpolation, which learns the relationship between consecutive traces to predict the intermediate trace along the dense inline receiver direction within common-shot gathers, ultimately bridging the gap in the coarse crossline receiver direction within common-shot gathers.

Recently, increasing attention has been given to self-supervised learning approaches, which mitigate the dependency on labeled crossline receiver direction within common-shot gathers. Chen *et al.*[39] proposed an interpolation method that combines POCS-Net, a CNN-based architecture built upon the POCS algorithm, with a self-supervised transfer learning framework. In addition, Wang *et al.*[40] presented a dip neural network (DINN) that leverages self-supervised learning for crossline interpolation. More recently, Goyes-Peñafiel *et al.*[41] introduced constrained diffusion-driven deep image prior (CDDIP), an unsupervised framework that combines diffusion processes with deep image prior (DIP) to reconstruct missing seismic traces without requiring labeled data.

In this study, we extend this strategy[39-41] for cross-wavefield reconstruction by utilizing not only 2D data from the inline receiver direction within common-shot gathers but also 3D volume data encompassing both inline and crossline receiver direction within common-shot gathers for training, effectively enhancing data interpolation by utilizing information across multiple spatial dimensions. The proposed approach, called the transposed arrangement strategy, alters the input array during the training and inference stages of the crossline interpolation model. This strategy allows for the reconstruction of the crossline receiver direction within common-shot gathers from original field data without the need for specific crossline labels. It involves training the model using a 3D input dataset from which 50% of the traces are regularly removed along the inline receiver direction within common-shot gathers. After training, this original input is transposed and fed into the model to facilitate data generation in the crossline receiver direction within common-shot gathers. In addition, to align event patterns between the inline and crossline receiver direction within common-shot gathers, linear moveout (LMO) correction is applied. Compared with existing self-supervised approaches, which rely on contrastive objectives, masked autoencoding, or physics-based regularization to avoid the need for labels, our method achieves self-supervision in a different way: It directly leverages dense inline traces as surrogate labels during training and then transposes the model for crossline inference. This provides a simple yet effective mechanism for eliminating the need for explicit crossline labels while remaining complementary to prior self-supervised paradigms. In addition, unlike other frameworks, such as DINN and POCS-Net, the proposed transposed strategy requires no auxiliary convex constraints, dip picking, or secondary network—the self-supervision emerges purely from a simple tensor transpose operation, making the method both conceptually elegant and computationally efficient. For this purpose, we employed two interpolation models: 3D U-Net and 3D U-Net+, whose effectiveness was measured against the conventional bicubic interpolation method,[2] the f–x interpolation method,[5] POCS[6] and CDDIP[41] through comparative performance evaluations. Among these, f–x interpolation was included as a widely used conventional method, POCS was selected as a representative compressive sensing–based method due to its simplicity, robustness, and computational efficiency, and CDDIP was considered to represent recent self-supervised deep learning approaches.

2. Methodology

2.1. Transposed arrangement strategy

Our proposed transposed arrangement strategy offers a label-substituting approach for interpolating actual field data. This strategy is universally applicable across deep learning models and is particularly optimized for scenarios where the receiver-to-streamer spacing ratio is 1–2. For example, in the dataset used in this study, the inline receiver spacing was 20 m and the crossline streamer spacing was 40 m, resulting in a ratio of 1:2. The proposed strategy unfolds as follows:

Initially, the original data D is acquired from the actual field and consists of N_i traces in the inline receiver direction within common-shot gathers, N_x traces in the crossline receiver direction within common-shot gathers, and N_t time samples. This can be represented as a matrix D with dimensions $N_x \times N_i \times N_t$, and the trace T can be illustrated along the time sample axis as follows:

$$\mathrm{D} = \begin{bmatrix} \mathrm{T}_{1,1} & \cdots & \mathrm{T}_{1,N_i} \\ \vdots & \ddots & \vdots \\ \mathrm{T}_{N_x,1} & \cdots & \mathrm{T}_{N_x,N_i} \end{bmatrix} \tag{I}$$

Where $T_{x,i}$ is a trace vector corresponding to the crossline index *x* and inline index *i*, each with a dimension of $N_t \times 1$. Here, N_x denotes the number of traces in the crossline direction, N_i the number of traces in the inline direction, and N_t the number of time samples per trace. Thus, each element $T_{x,i}$ represents a seismic trace consisting of N_t amplitude values along the time axis.

From the original data *D*, a volume patch *P* of size $\tilde{N}_x \times \tilde{N}_i \times \tilde{N}_t$ is extracted. If the starting position of the patch is denoted as (*i*, *j*), then the patch *P* can be represented as:

$$P = \begin{bmatrix} T_{i,j} & \cdots & T_{i,j+\tilde{N}_i-1} \\ \vdots & \ddots & \vdots \\ T_{i+\tilde{N}_x-1,j} & \cdots & T_{i+\tilde{N}_x-1,j+\tilde{N}_i-1} \end{bmatrix} \tag{II}$$

Where $\tilde{N}_x$ and $\tilde{N}_i$ serve as the indices navigating the crossline and inline receiver direction within common-shot gathers, respectively, within the patch, and $\tilde{N}_i = 2\tilde{N}_x - 1$.

During the training phase, the complete patch *P* is used as a label, while an input P_{in} is created by regularly removing 50% of P in the inline receiver direction within common-shot gathers, represented as follows:

$$P_{in} = \begin{bmatrix} T_{i,j} & 0 & T_{i,j+2} & 0 & \cdots & T_{i,j+\tilde{N}_i-1} \\ T_{i+1,j} & 0 & T_{i+1,j+2} & 0 & \cdots & T_{i+1,j+\tilde{N}_i-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ T_{i+\tilde{N}_x-1,j} & 0 & T_{i+m-1,j+2} & 0 & \cdots & T_{i+\tilde{N}_x-1,j+\tilde{N}_i-1} \end{bmatrix} \tag{III}$$

The training process for the reconstruction model can be represented by the following equation:

$$\hat{P} = Net\left(P_{in}\right)$$

$$\min imize\, L(\hat{P}, P) \tag{IV}$$

Where *Net*(.) denotes the reconstruction deep learning model, $\hat{p}$ is the prediction of the model, *L*(.)is the loss function used to train the model. This approach trains the reconstruction model to fill in the missing traces represented by the zero values in **Equation III.**

In the inference phase, a new input patch is created by: (i) Extracting a patch of size $\tilde{N}_x \times \tilde{N}_x$ with $\tilde{N}_t$ time sampling from the original data *D*, (ii) regularly adding zero values into the crossline receiver direction within common-shot gathers expanding the dimensions to $(2\tilde{N}_x - 1) \times \tilde{N}_x \times \tilde{N}_t$, and (iii) transposing the first and second dimensions to result in a final dimension of $\tilde{N}_x \times (2\tilde{N}_x - 1) \times \tilde{N}_t$. Then, the new input patch for reconstructing crossline receiver direction within common-shot gathers can be represented as follows:

$$\check{P}_{in} = \begin{bmatrix} T_{i,j} & T_{i,j+2} & \cdots & T_{i,j+2(\tilde{N}_x-1)} \\ 0 & 0 & \cdots & 0 \\ T_{i+1,j} & T_{i+1,j+2} & \cdots & T_{i+1,j+2(\tilde{N}_x-1)} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ T_{i+\tilde{N}_x-1,j} & T_{i+\tilde{N}_x-1,j+2} & \cdots & T_{i+\tilde{N}_x-1,j+2(\tilde{N}_x-1)} \end{bmatrix}^T \tag{V}$$

Subsequently, by inputting new input patches $\check{P}_{in}$ from the original data D into the trained model, Net(.), the model is able to generate the crossline receiver direction within common-shot gathers between streamers. Figure 1 visually illustrates examples for **Equations III** and **V**.

2.2. LMO correction

In the towed streamer system, the difference in offsets between the source and receivers generates varying moveout patterns in the inline and crossline receiver direction within common-shot gathers. Specifically, the moveouts in the inline receiver direction within common-shot gathers are steeper compared to those in the crossline. To harmonize these pattern discrepancies between the inline and crossline receiver direction within common-shot gathers, LMO correction is employed.[38]

LMO correction is a data processing technique used to adjust the arrival time differences of seismic waves captured by multiple receivers. This method is primarily used to improve the initial alignment of multi-channel data and to reduce the time differences between records at close ranges. Notably, this correction can be applied using just the streamer and receiver intervals, without the need for actual subsurface velocity information.[42] The formula for the LMO correction used in this study is as follows:

$$\tau = t\text{-}px \tag{VI}$$

Where *t* represents the original time of the seismic event at offset *x*, *p* is the ray parameter defined as *dt*/*dx*, and *τ* is the time-shift for the corrected events. In this study, LMO correction is applied to the original seismic data before fetching the volume patches, ensuring that event patterns are similar in both inline and crossline receiver direction within common-shot gathers, with a typical dip range reduced from approximately −8 ms/trace before correction to within ±2 ms/trace after correction. The ray parameter was estimated from the first arrivals and applied uniformly across the line for processing simplicity. In practice, this step could be replaced or refined using spatially varying normal moveout correction or pre-stack time migration velocity fields to achieve higher physical realism.

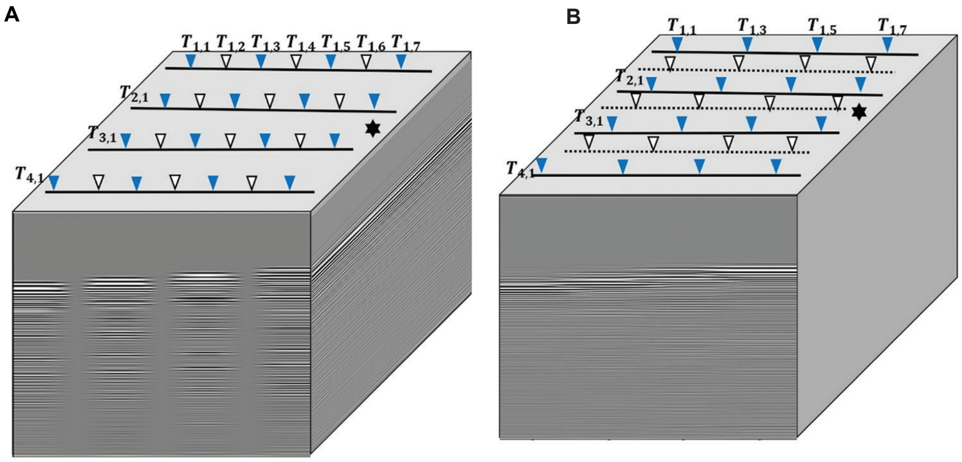


Figure 1. Input arrays for the transposed arrangement strategy during training (A) and inference stages (B). Blue triangles indicate active receivers, and white triangles denote absent receivers, represented as zeros in the data array. Black solid lines represent streamers, while dotted lines show the positions of absent streamers.

2.3. 3D U-Net and 3D U-Net+

Given that our proposed strategy can be applied to any deep learning model, we employed the widely used 3D U-Net and its extension, 3D U-Net+, as our interpolation models.[43,44] Both models were implemented using 3×3×3 convolution layers and 1×1×2 upsampling layers. The 3D U-Net utilizes rectified linear unit as its activation function, whereas 3D U-Net+ employs exponential linear unit. During the model training phase, we adopted the mean absolute error as the loss function and the Adam optimizer for optimization. The learning rate was set to 0.0001. Figure 2 illustrates the structures of 3D U-Net and 3D U-Net+. In this study, the "+" in 3D U-Net+ denotes an extension of the basic 3D U-Net by adding skip connections, which strengthens feature propagation and better preserves structural information during interpolation.

3. Data example

We applied our transposed arrangement strategy to the publicly available SEG/EAGE Salt Model, specifically to the narrow azimuth data. The acquisition parameters for this dataset are detailed in Table 1.[45] The receiver and streamer spacings are 20 m and 40 m, respectively, maintaining a 1:2 ratio, which is well-suited for the application of our strategy.

To evaluate the performance of our proposed strategy, we initially removed 50% of the data regularly in both inline and crossline receiver direction within common-shot gathers from the original dataset. Consequently, the number of streamers was adjusted from 8 to 4, the receiver spacing from 20 m to 40 m, and the streamer spacing from 40 m to 80 m. The removed crossline receiver direction within common-shot gathers were then utilized as the ground truth in our final model evaluation, and to distinguish this dataset from the test set used to train the model below, we call it the crossline test dataset.

Using the adjusted dataset, we initially applied LMO correction based on the first arrival to mitigate discrepancies of moveout patterns arising from inline and crossline receiver direction within common shot gathers, as illustrated in Figure 3. Then, we split the dataset into training, validation, and test sets in a 7:2:1 ratio, respectively, according to shot locations.

During the training phase, we extracted volume patches of size 4×7×624 from the training and validation datasets. To generate a sufficient number of patches, each patch was configured to include seven receivers. This setup ensured that there were overlapping sections between patches. Each patch consists of 4×7 traces, aligned along the crossline and inline receiver direction within common-shot gathers, respectively, with each trace containing 624 time samples. These complete patches are used for labels. Then, we regularly removed 50% of each volume patch in the inline receiver direction within common-shot gathers and replaced the removed values with zeros, to serve as input for our deep learning models as shown in Figure 1A. Thus, the optimal patch size (4×7) for our strategy is determined based on the number of streamers and the number of receivers per streamer.

We trained models using both 3D U-Net and 3D U-Net+ architectures. After training, we evaluated their performance on the test dataset. As illustrated in Figure 4, both models demonstrated similar residual levels and provided interpolation results that closely matched the labels. This performance assessment confirms that our

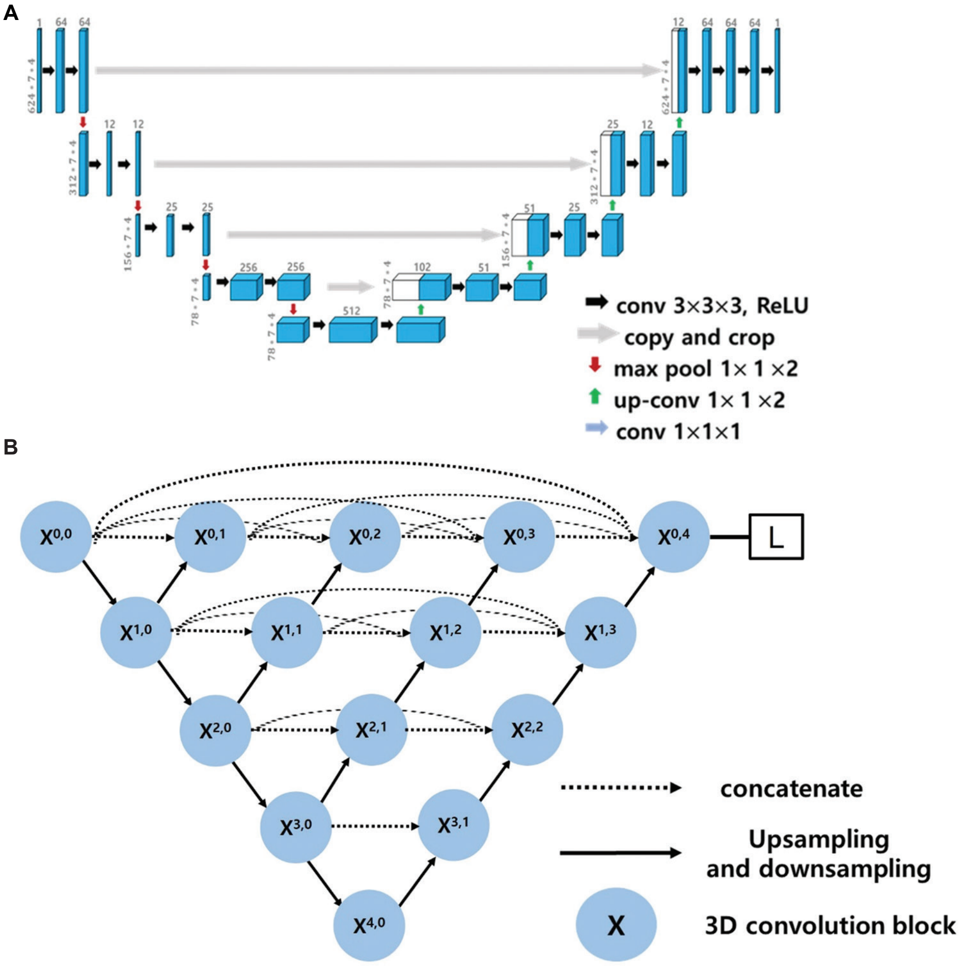


Figure 2. The architectures of 3D U-Net (A) and 3D U-Net+ (B)

Table 1. Data acquisition parameter

Acquisition geometry	Value
Number of sail lines	50
Spacing of sail lines (m)	160
Shots per line	96
Spacing of shot (m)	80
Number of streamers	8
Spacing of receiver (m)	20
Receivers per streamer	68
Max offset (m)	1,340
Number of samples	625
Spacing of samples (s)	0.008
Spacing of streamer (m)	40

models are effectively capable of interpolating data, with 50% of dataset removed regularly in the inline receiver direction within common-shot gathers.

Finally, to reconstruct the crossline receiver direction within common-shot gathers, we implemented the transposed arrangement strategy by feeding transposed input arrays into our trained models, and subsequently compared the outputs with the ground truth from the crossline test dataset. As depicted in Figure 5, the 3D U-Net+ model exhibited smaller residuals compared to the 3D U-Net, POCS, CDDIP, the f–x interpolation method, and the bicubic interpolation method, with the latter showing significantly larger discrepancies relative to the deep learning models. In addition, Figure 5H and 5O represent the results of training and testing the 3D U-Net+ on original data without applying LMO. These results demonstrate that applying LMO significantly improved the performance of the model.

To further highlight differences in fine-scale textures, Figure 6 presents enlarged views of the low amplitude regions marked by red boxes in Figure 5. The results show that the proposed 3D U-Net+ achieves more accurate

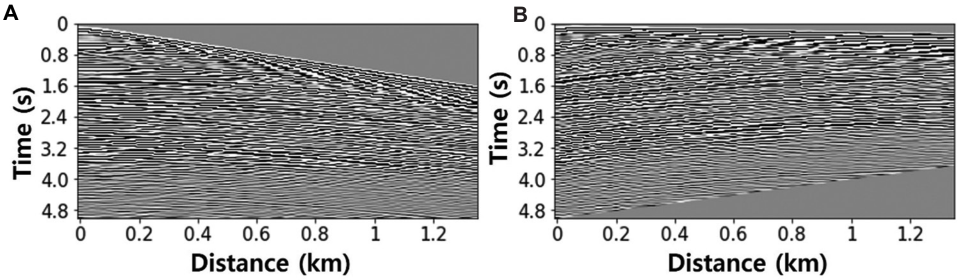


Figure 3. Inline receiver direction within common-shot gathers with before (A) and after applying linear moveout correction (B)

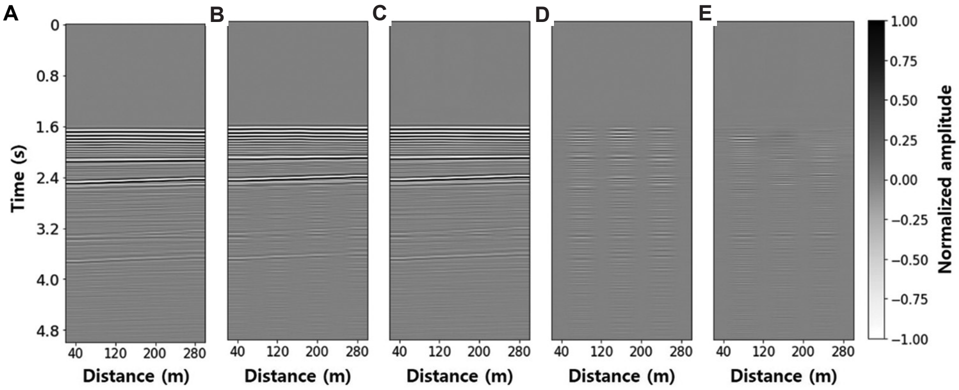


Figure 4. Interpolation results for the inline receiver direction within common shot gathers. (A) Ground truth. (B) 3D U-Net. (C) 3D U-Net+. (D) Difference between (A) and (B). (E) Difference between (A) and (C).

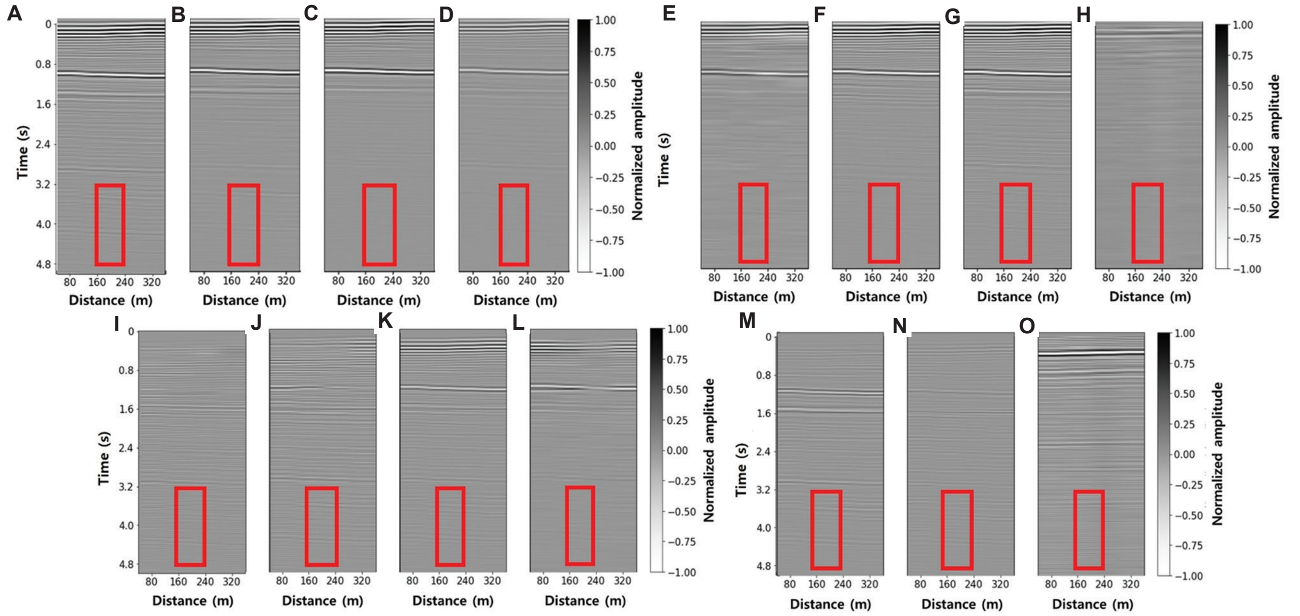


Figure 5. Linear moveout (LMO)–corrected interpolation results for crossline receiver direction within common-shot gathers. (A) Ground truth. (B) Bicubic. (C) f–x interpolation. (D) Projection onto convex sets. (E) Constrained diffusion-driven deep image prior. (F) 3D U-Net. (G) 3D U-Net+. (H) 3D U-Net+ (LMO not applied). (I) Difference between (A) and B). (J) Difference between (A) and (C). (K) Difference between (A) and (D). (L) Difference between (A) and (E). (M) Difference between (A) and (F). (N) Difference between (A) and (G). (O) Difference between (F) and ground truth without LMO applied.

recovery of subtle structures and produces smaller residuals compared to conventional methods, even in low amplitude zones.

Further analysis of trace comparisons, as illustrated in Figure 7, indicates that both 3D U-Net and 3D U-Net+ accurately predict large amplitude events with flat slopes

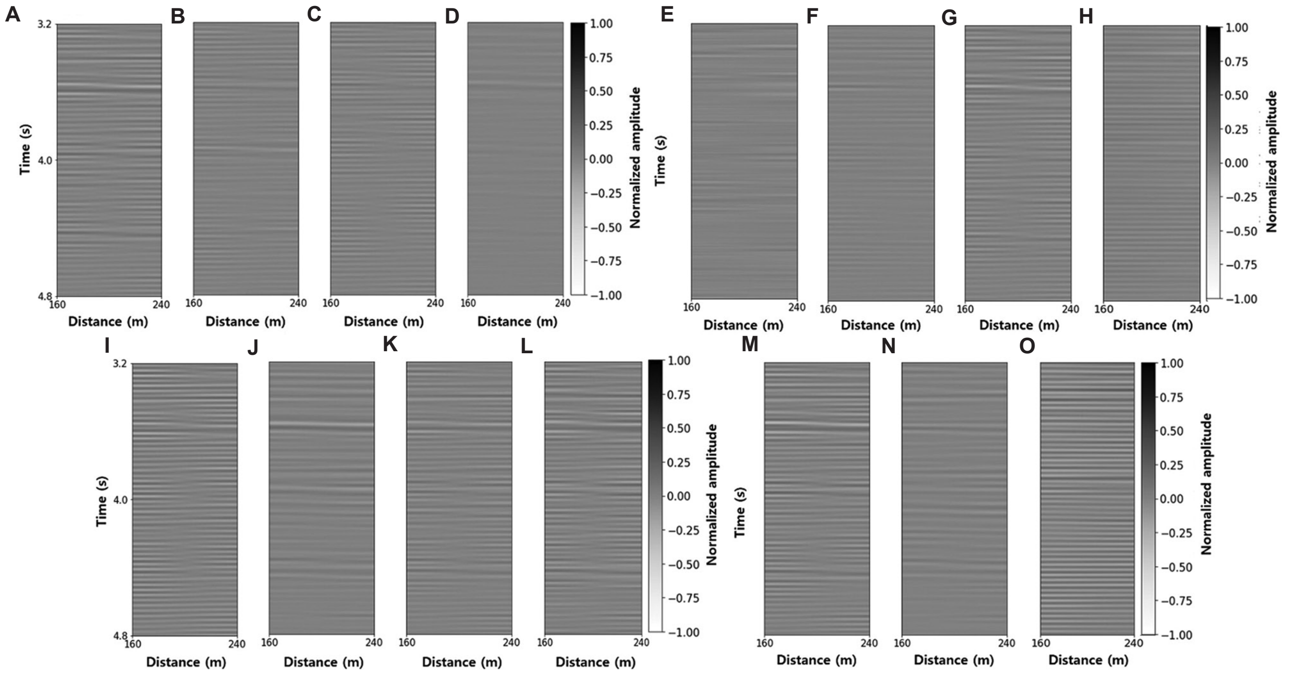


Figure 6. Enlarged views of the regions indicated by red boxes in Figure 5, focusing on the weak-amplitude zones. (A) Ground truth. (B) Bicubic. (C) f–x interpolation. (D) Projection onto convex sets. (E) Constrained diffusion-driven deep image prior. (F) 3D U-Net. (G) 3D U-Net+. (H) 3D U-Net+ (linear moveout [LMO] not applied). (I) Difference between (A) and (B). (J) Difference between (A) and (C). (K) Difference between (A) and (D). (L) Difference between (A) and (E). (M) Difference between (A) and (F). (N) Difference between (A) and (G). (O) Difference between (F) and ground truth without LMO applied.

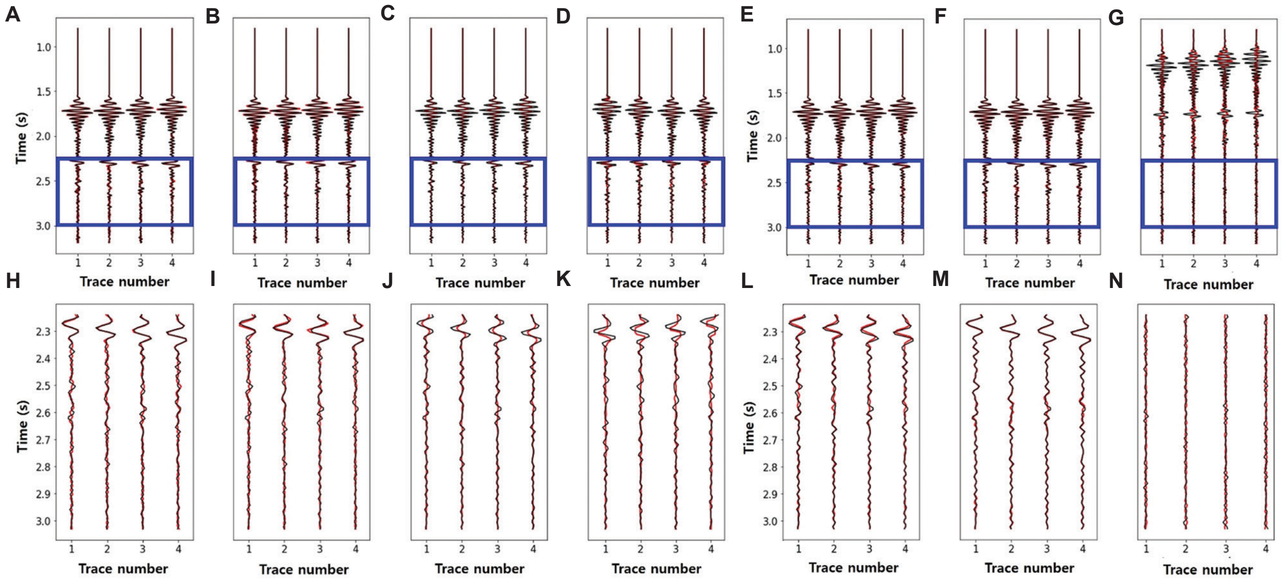


Figure 7. Comparison of traces for Figure 5. The black line represents the trace for the ground truth, and the red line represents the trace for the prediction result. (A) Bicubic. (B) f–x interpolation. (C) Projection onto convex sets. (D) Constrained diffusion-driven deep image prior. (E) 3D U-Net. (F) 3D U-Net+. (G) 3D U-Net+ (linear moveout [LMO] not applied). (H–N) Enlarged views of the blue box in (A), (B), (C), (D), (E), (F), and (G), respectively.

and also achieve high precision in reconstructing small amplitude events with steep slopes. In contrast, both the bicubic method, f–x interpolation method and POCS method struggle with the interpolation of small amplitude events with steep slopes. CDDIP is a self-supervised interpolation method whose performance is highly sensitive to hyperparameter choices, which can substantially affect the quality of reconstructed results. In our experiments, we adopted the publicly available code with default hyperparameters. As shown in Figure 7, while CDDIP successfully predicted the overall seismic signals, its reconstruction accuracy in fine-scale details was relatively low compared to our proposed method.

Table 2. PSNR and SSIM results

Method	PSNR (standard deviation)	SSIM (standard deviation)
bicubic	30.35 (3.93)	0.60 (0.07)
f–x interpolation	30.79 (3.22)	0.61 (0.11)
POCS	31.21 (3.56)	0.62 (0.14)
CDDIP	31.59 (4.08)	0.68 (0.22)
3D U-Net	32.23 (1.11)	0.71 (0.21)
3D U-Net+	32.86 (3.65)	0.74 (0.11)
3D U-Net+ (without LMO)	20.36 (1.71)	0.44 (0.01)

Abbreviations: CDDIP: Constrained diffusion-driven deep image prior; LMO: Linear moveout; POCS: Projection onto convex sets; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index.

For quantitative performance evaluation, we used peak signal-to-noise ratio (PSNR) as performance metrics, which are calculated using following equations:

$$PSNR = 10\log_{10}(peakval^2)/MSE \tag{VII}$$

Where peakval (peak value) is the maximum value in the data. A larger PSNR indicates better quality of the data.

Table 2 presents the PSNR results for the test dataset, with the 3D U-Net+ model demonstrating higher value of 32.86 dB, calculated using data normalized to a unit peak amplitude (max = 1). Consequently, it has been demonstrated that, even in the absence of label data for actual field data, high-accuracy interpolation can be

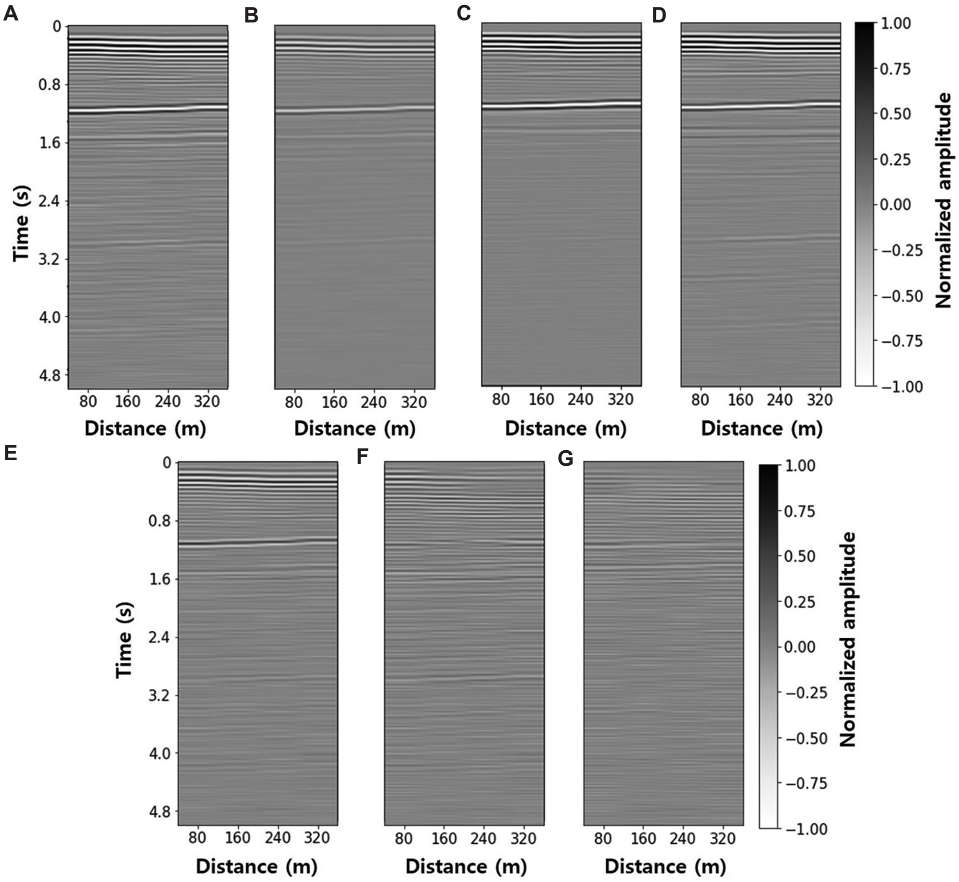


Figure 8. Interpolation results for crossline receiver direction within common-shot gathers with 5% Gaussian noise. (A) Ground truth. (B) Constrained diffusion-driven deep image prior (C) 3D U-Net. (D) 3D U-Net+. (E) Difference between (A) and (B). (F) Difference between (A) and (C). (G) Difference between (A) and (D).

performed through the transposed arrangement strategy-based deep learning interpolation method.

Figure 8 presents additional experiments conducted by adding 5% Gaussian noise to the original test data and evaluating the interpolation performance of CDDIP, 3D U-Net, and 3D U-Net+, which previously demonstrated relatively higher accuracy. The results show that all three methods were able to effectively handle a certain level of noise, but the proposed 3D U-Net+ produced smaller residuals compared to both CDDIP and 3D U-Net, indicating superior robustness in noisy conditions. Nevertheless, to obtain even better results, additional denoising would be necessary.

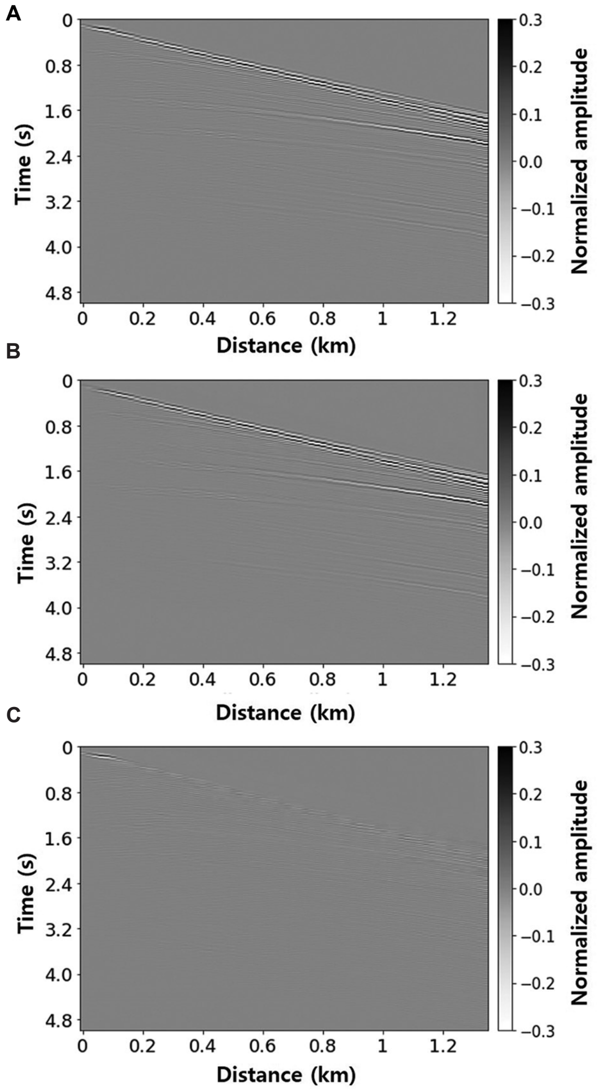


Figure 9. Inline receiver direction within common-shot gathers with the interpolated crossline location of 3D U-Net+. (A) Ground-truth. (B) Interpolation result. (C) Difference between (A) and (B).

Table 3 presents a comparison of the computational costs of the proposed 3D U-Net and 3D U-Net+ models against conventional interpolation methods, including bicubic, f–x interpolation, POCS, and CDDIP. The experiments were conducted on a server equipped with four NVIDIA TITAN RTX GPUs (each with 24 GB memory), running CUDA version 12.4 and driver version 550.76. The reported memory usage was measured as the peak GPU/CPU memory consumption during training and inference. The table summarizes training time (GPU hours), inference time, and GPU/CPU memory usage for each method.

Figure 9 shows the inline receiver direction within common-shot gathers with the reconstructed crossline position, combined with the output patches from the 3D U-Net+ model, which is identified as our optimal model. The results demonstrate that the proposed model is able to accurately predict the overall seismic signals, effectively capturing both strong reflections and subtle structures. Figure 10 presents the f–x spectrum corresponding to Figure 9, illustrating the frequency–wavenumber characteristics of the reconstructed seismic section. Although certain high-frequency components are not fully recovered, the overall spectrum closely resembles that of the label data, indicating that the proposed method preserves the dominant frequency content and structural consistency. Figure 11 presents a further trace-based analysis of the results in Figure 9 to assess the accuracy of the interpolation at the individual trace level.

4. Discussion

Our proposed method is designed for cross-wavefield reconstruction in marine towed-streamer systems with regularly sampled data and a 1:2 spacing ratio between receivers and streamers. The experiment utilizes complete

Table 3. Computational cost comparison

Method	Training time (GPU hours)	Inference time (min)	GPU/CPU memory usage
Bicubic	-	~0.1	<1 GB
f–x interpolation		~0.5	<1 GB
POCS		~10	<2 GB
CDDIP	~200	~30	~6 GB (training)/<2 GB (inference)
3D U-Net	~120	~2	~23 GB (training)/<1 GB (inference)
3D U-Net+	~150	~2	~24 GB (training)/<1 GB (inference)

Abbreviations: CDDIP: Constrained diffusion-driven deep image prior; POCS: Projection onto convex sets.

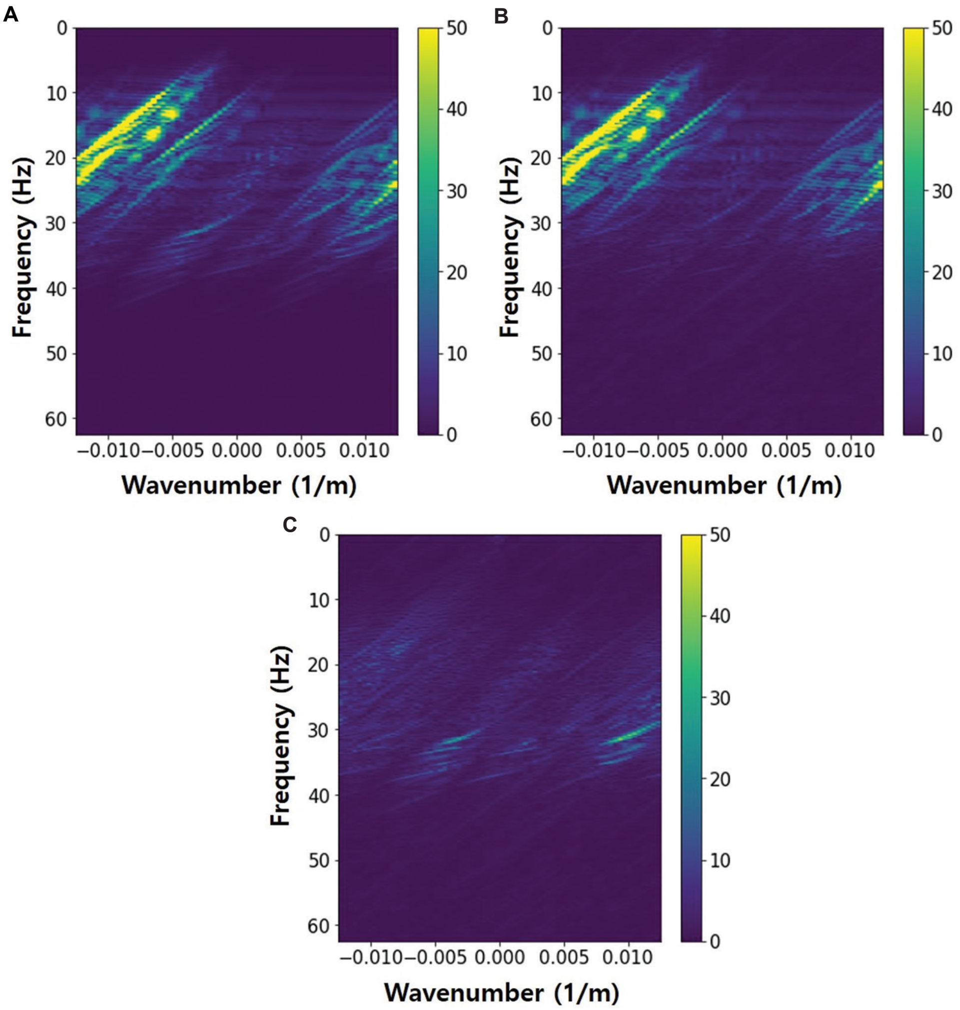


Figure 10. f–x spectrum corresponding to Figure 9. (A) Ground-truth. (B) Interpolation result. (C) Difference between (A) and (B).

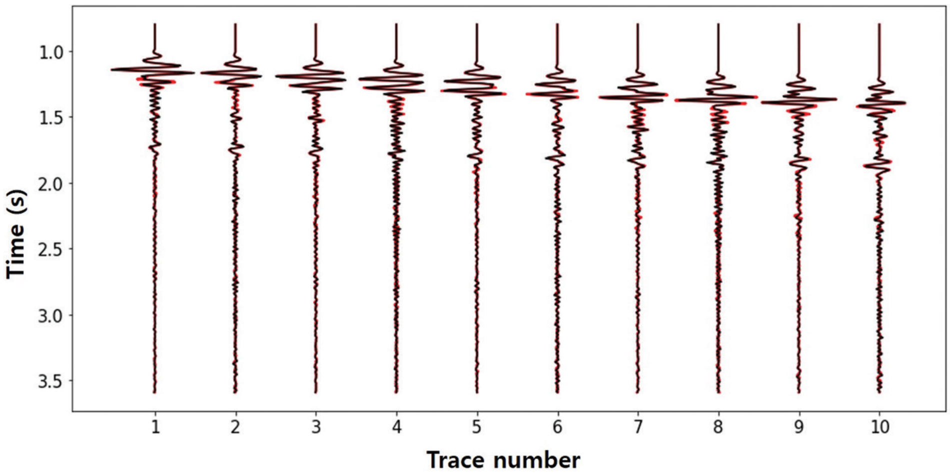


Figure 11. Comparison of traces 40–49 (out of 624) extracted from the gathers shown in Figure 9. The black line represents the trace for the ground truth, and the red line represents the trace for the prediction result.

common-shot gathers from towed-streamer data; however, regularization may be necessary if missing data is present. In addition, our method can be extended to CMP gathers or seismic image volumes, provided that the data maintain a regular 1:2 spacing ratio. If the spacing ratio is 2^n, where n >1, our method can be applied recursively to interpolate the data to a 1:1 ratio. If the spacing ratio is not in the form of $1:2^n$, the transposed arrangement strategy may be difficult to apply, and additional data processing may be necessary to adjust to a $1:2^n$ ratio. However, it should be noted that recursive applications may accumulate errors, potentially impacting accuracy.

In addition, the reconstruction performance of the proposed strategy is highly dependent on the quality of inline receiver direction within common-shot gathers, since dense inline traces are employed as surrogate labels during training. If the inline receiver direction within common-shot gathers are noisy or contain structural discontinuities, the errors may propagate into the crossline reconstruction. The method may also be sensitive to strong random or coherent noise, which could reduce accuracy in weak-amplitude regions. Future improvements could incorporate denoising pre-processing or noise-aware loss functions to enhance robustness.

The experiments were conducted on a server equipped with four NVIDIA TITAN RTX GPUs (each with 24 GB memory), running CUDA version 12.4 and driver version 550.76. During training, each model instance occupied approximately 23 GB of GPU memory per device, nearly exhausting the available capacity of a single TITAN RTX card. This high memory demand arises mainly from the use of 3D convolutional kernels and large input patches (*e.g.*, 4 × 7 × 624 traces). Therefore, multi-GPU environments with sufficient memory capacity (≥24 GB per GPU) are strongly recommended for training 3D U-Net and U-Net+ architectures with the proposed transposed arrangement strategy. In this study, the batch size during training was set to 128, and we note that reducing the batch size would alleviate the memory issue. In contrast, the inference stage required <1 GB of GPU memory, making the trained models practical for deployment in real-world seismic data processing tasks.

The seismic dataset used in this study consists of 50 sail lines, each containing 96 shots, 8 streamers, and 68 receivers per streamer. Each receiver records 625 samples at an 8 ms sampling interval, resulting in a total data volume of approximately 6.1 GB for a single-component (float32) dataset.

The proposed 3D U-Net+ model required about 24 GB of GPU memory per device during training and was trained using four TITAN RTX GPUs for approximately 150 GPU h. In contrast, the inference phase required <1 GB of GPU memory, demonstrating that the proposed framework can process seismic data efficiently even at moderate survey scales. Thus, although model training requires high-performance GPUs (≥24 GB each), the trained network can be efficiently applied to large-scale seismic data in practical production environments. This method was developed for streamer data on regular grids and is not directly applicable to irregular land geometries.

5. Conclusion

This study introduced the transposed arrangement strategy, a novel approach designed to enhance the performance of crossline interpolation. By adjusting input array configurations during the training and inference phases, our method successfully reconstructed crossline receiver direction within common-shot gathers without the need for labeled data in the crossline receiver direction within common-shot gathers. To address feature differences between inline and crossline receiver direction within common-shot gathers, we applied LMO correction, enabling effective crossline interpolation using a model trained with inline receiver direction within common-shot gathers.

The interpolation performance of our method was rigorously evaluated by comparing it with traditional bicubic interpolation methods, f–x interpolation method, and POCS using deep learning models, 3D U-Net and 3D U-Net+. The results demonstrated that 3D U-Net+ provided the clearest and most accurate interpolations, closely resembling the original data. Numerically, 3D U-Net+ also exhibited superior PSNR values, confirming its efficacy.

This research confirms that applying the proposed method in deep learning-based models significantly enhances the accuracy of crossline interpolation, suggesting a promising direction for future advancements in seismic data processing.

Acknowledgments

None.

Funding

This work was supported by the Korea Institute of Marine Science and Technology Promotion grant funded by the Ministry of Oceans and Fisheries (20220254, Development of technology for seabed classification based on machine learning).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Yonghwan Joo, Daeung Yoon
Formal analysis: Jiyun Yu
Investigation: Jiyun Yu
Methodology: Yonghwan Joo, Daeung Yoon
Writing–original draft: Jiyun Yu
Writing–review & editing: Jiyun Yu, Daeung Yoon

Availability of data

The data underlying this article are available in the SEG wiki (https://wiki.seg.org/wiki/SEG_C3_NA).

Further disclosure

Part of the findings of this study were presented at the 2023 Fall Joint Conference and the 9th International Symposium on Mine Reclamation, held at Pyeongchang on November 03, 2023.

References

1. Ikelle LT, Amundsen L. *Introduction to Petroleum Seismology*. Society of Exploration Geophysicists; 2005.

   doi: 10.1190/1.9781560803447.ch1

2. Keys R. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust.* 1981;29(6):1153-1160.

   doi: 10.1109/TASSP.1981.1163711

3. Ronen J. Wave-equation trace interpolation. *Geophysics*. 1987;52(7):973-984.

   doi: 10.1190/1.1442366

4. Fomel S. Seismic reflection data interpolation with differential offset and shot continuation. *Geophysics*. 2003;68(2):733-744.

   doi: 10.1190/1.1567243

5. Spitz S. Seismic trace interpolation in the F-X domain. *Geophysics*. 1991;56(6):785-794.

   doi: 10.1190/1.1443096

6. Spitz S. Pattern recognition, spatial predictability, and subtraction of multiple events. *Geophysics*. 1999;64(2):431-436.

   doi: 10.1190/1.1438154

7. Bregman L. The method of successive projection for finding a common point of convex sets (theorems for determining common point of convex sets by method of successive projection). *Sov Math*. 1965;6:688-692.

8. Herrmann FJ, Hennenfent G. Non-parametric seismic data recovery with curvelet frames. *Geophys J Int*. 2008;173(1):233-248.

   doi: 10.1111/j.1365-246X.2007.03698.x

9. Naghizadeh M, Sacchi MD. Multistep autoregressive reconstruction of seismic records. *Geophysics*. 2007;72(6):V111-V118.

   doi: 10.1190/1.2771685

10. Trickett S, Burroughs L, Milton A, Walton L, Dack R. Rank-Reduction-Based Trace Interpolation. In: *Conference: SEG Technical Program Expanded Abstracts 2010*; 2010. p. 3829-3833.

    doi: 10.1190/1.3513645

11. Gao J, Sacchi MD, Chen X. A fast reduced rank interpolation method for prestack seismic volumes that depend on four spatial dimensions. *Geophysics*. 2013;78(1):V21-V30.

    doi: 10.1190/geo2012-0038.1

12. Ma J. Three-dimensional irregular seismic data reconstruction via low-rank matrix completion. *Geophys*. 2013;78(5):V181-V192.

    doi: 10.1190/geo2012-0465.1

13. Abma R, Kabir N. 3D interpolation of irregular data with a POCS algorithm. *Geophysics*. 2006;71(6):E91-E97.

    doi: 10.1190/1.2356088

14. Gan S, Wang S, Chen Y, Zhang Y, Jin Z. Dealiased seismic data interpolation using seislet transform with low-frequency constraint. *IEEE Geosci Remote Sens Lett.* 2015;12(10):2150-2154.

    doi: 10.1109/LGRS.2015.2453119

15. Liu W, Cao S, Gan S, Chen Y, Zu S, Jin Z. One-step slope estimation for dealiased seismic data reconstruction via iterative seislet thresholding. *IEEE Geosci Remote Sens Lett.* 2016;13(10):1462-1466.

    doi: 10.1109/LGRS.2016.2591939

16. Jia Y, Ma J. What can machine learning do for seismic data processing? An interpolation application. *Geophysics*. 2017;82:163-177.

    doi: 10.1190/geo2016-0300.1

17. Oliveira DAB, Ferreira RS, Silva R, Brazil EV. Interpolating seismic data with conditional generative adversarial networks. *IEEE Geosci Remote Sens Lett*. 2018;15:1952-1956.

    doi: 10.1109/LGRS.2018.2866199

18. Wang B, Zhang N, Lu W, Geng J, Huang X. Intelligent missing shots' reconstruction using the spatial reciprocity of green's function based on deep learning. *IEEE Trans Geosci Remote Sens*. 2020;58:1587-1597.

    doi: 10.1109/TGRS.2019.2947085

19. Park J, Choi J, Seol SJ, Byun J, Kim Y. A method for adequate selection of training data sets to reconstruct seismic data using a convolutional U-Net. *Geophysics*. 2021;86:375-388.

doi: 10.1190/geo2019-0708.1

20. Park J, Yeeh Z, Seol SJ, *et al*. Seismic data interpolation using attention-based deep learning. *Eur Assoc Geosci Eng*. 2022;2022:1-5.

doi: 10.3997/2214-4609.202210245

21. Yu J, Wu B. Attention and hybrid loss guided deep learning for consecutively missing seismic data reconstruction. *IEEE Trans Geosci Remote Sens*. 2021;60:1-8.

doi: 10.1109/TGRS.2021.3068279

22. Li X, Wu B, Zhu X, Yang H. Consecutively missing seismic data interpolation based on coordinate attention unet. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/LGRS.2021.3128511

23. Lou Y, Wu L, Liu L, *et al*. Irregularly sampled seismic data interpolation via wavelet-based convolutional block attention deep learning. *Artif Intellig Geosci*. 2022;3:192-202.

doi: 10.1016/j.aiig.2022.12.001

24. Dodda VC, Kuruguntla L, Mandpura AK, Elumalai K. Simultaneous seismic data denoising and reconstruction with attention-based wavelet-convolutional neural network. *IEEE Trans Geosci Remote Sens*. 2023;61:1-14.

doi: 10.1109/TGRS.2023.3267037

25. Tian Y, Fu L, Fang W, Li T. FR-UNet: A feature restoration-based UNet for seismic data consecutively missing trace interpolation. *IEEE Trans Geosci Remote Sens*. 2025;63:1-10.

doi: 10.1109/TGRS.2025.3531934

26. Kaur H, Pham N, Fomel S. Seismic data interpolation using CycleGAN. *SEG Tech Prog Expand Abstracts*. 2019;38:5407.

doi: 10.1190/segam2019-3207424.1

27. Chang DK, Yang WY, Yong XS, Li S. Seismic data interpolation with conditional generative adversarial network in time and frequency domain. *SEG Tech Prog Expand Abstracts*. 2019;38:2589-2593.

doi: 10.1190/segam2019-3210118.1

28. Deng F, Wang S, Wang X, Fang P. Seismic data reconstruction based on conditional constraint diffusion model. *IEEE Geosci Remote Sens Lett*. 2024;21:1-5.

doi: 10.1109/LGRS.2024.3371675

29. Kong F, Picetti F, Lipari V, Bestagini P, Tang X, Tubaro S. Deep prior-based unsupervised reconstruction of irregularly sampled seismic data. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/LGRS.2020.3044455

30. Jin Z, Li X, Yang H, Wu B, Zhu X. Depthwise separable convolution Unet for 3D seismic data interpolation. *Front Earth Sci*. 2023;10:1-18.

doi: 10.3389/feart.2022.1005505

31. Chang D, Yang W, Young X, *et al*. Seismic data interpolation using dual-domain conditional generative adversarial networks. *IEEE Geosci Remote Sens Lett*. 2020;18:1856-1860.

doi: 10.1109/LGRS.2020.3008478

32. Dou Y, Li K, Duan H, *et al*. MDA GAN: Adversarial-learning-based 3-D seismic data interpolation and reconstruction for complex missing. *IEEE Trans Geosci Remote Sens*. 2023; 61:1-14.

doi: 10.48550/arXiv.2204.03197

33. Yu J, Yoon D. Crossline reconstruction of 3D seismic data using 3D cWGAN: A comparative study on sleipner seismic survey data. *Appl Sci*. 2023;13:5999.

doi: 10.3390/app13105999

34. Ding M, Zhou Y, Chi Y. Self-attention generative adversarial network interpolating and denoising seismic signals simultaneously. *Remote Sens*. 2024;16:305.

doi: 10.3390/rs16020305

35. Wang S, Deng F, Jiang P, Gong Z, Wei X, Wang Y. SeisFusion: Constrained diffusion model with input guidance for 3-D seismic data interpolation and reconstruction. *IEEE Trans Geosci Remote Sens*. 2024;62:1-15.

doi: 10.48550/arXiv.2403.11482

36. Qian F, Liu Z, Wang Y, Liao S, Pan S, Hu G. DTAE: Deep tensor autoencoder for 3-D seismic data interpolation. *IEEE Trans Geosci Remote Sens*. 2021;60:1-19.

doi: 10.1109/TGRS.2021.3075968

37. Larsen Greiner TA, Hlebnikov V, Lie JE, *et al*. Cross-streamer wavefield reconstruction through wavelet domain learning. *Geophysics*. 2020;85:457-471.

doi: 10.1190/geo2019-0771.1

38. Yeeh Z, Yoon D, Byun J. Cross-streamer wavefield reconstruction of a towed streamer system using bidirectional LSTM networks with a traces-to-trace approach. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

doi: 10.1109/LGRS.2023.3312712

39. Chen Y, Yu S, Lin R. Self-supervised transfer learning POCS-net for seismic data interpolation. *IEEE Trans Geosci Remote Sens*. 2024;62:1-18.

doi: 10.1109/TGRS.2024.3494723

40. Wang S, Wu X, Chen J. Dip-informed neural network for self-supervised anti-aliasing seismic data interpolation. *IEEE Trans Geosci Remote Sens*. 2024;62:1-4.

doi: 10.1109/TGRS.2024.3359247

41. Goyes-Peñafiel P, Kamilov US, Arguello H. CDDIP:

Constrain diffusion-driven deep image prior for seismic data reconstruction. *IEEE Geosci Remote Sens Lett.* 2025;22:1-5.

doi: 10.48550/arXiv.2407.17402

42. Souza WE, Cerqueira AG, Porsani MJ. First-break prediction in 3-D land seismic data using the dynamic time warping algorithm. *Geophys J Int*. 2024;237:402-418.

    doi: 10.1093/gji/ggae048

43. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Cham: Springer; 2015. p. 234-241.

44. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A nested U-Net architecture for medical image segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support*. 2018;11045:3-11.

    doi: 10.48550/arXiv.1807.10165

45. Ji J, Choi YG. 3D seismic data processing methodology using public domain software system. *Geophys Geophys Explor.* 2010;13:159-168.

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Hybrid convolutional neural network–graph attention network–gradient boosting decision tree model for seismic impedance inversion prediction

Tianwen Zhao[1], Guoqing Chen[2], Cong Pang[3,4], Palakorn Seenoi[5], Nipada Papukdee[6], Piyapatr Busababodhin[7], and Yiru Du[2]*

[1]Department of Trade and Logistics, Daegu Catholic University, Gyeongsan, Daegu, Republic of Korea

[2]Mathematical Modeling Research Center, Chengdu Jincheng College, Chengdu, Sichuan, China

[3]Institute of Seismology, China Earthquake Administration, Wuhan, Hubei, China

[4]Wuhan Gravitation and Solid Earth Tides, National Observation and Research Station, Wuhan, Hubei, China

[5]Department of Statistics, Faculty of Science, Khon Kaen University, Mueang Khon Kaen, Khon Kaen, Thailand

[6]Department of Applied Statistics, Rajamangala University of Technology, Isan Khon Kaen Campus, Mueang Khon Kaen, Khon Kaen, Thailand

[7]Department of Mathematics, Faculty of Science, Mahasarakham University, Kantharawichai, Maha Sarakham, Thailand

***Corresponding author:**
Yiru Du
(duyiru@cdjcc.edu.cn)

Citation: Zhao T, Chen G, Pang C, *et al.* Hybrid convolutional neural network–graph attention network–gradient boosting decision tree model for seismic impedance inversion prediction. *J Seismic Explor*. 2025;34(5):81-98. doi: 10.36922/JSE025310051

Received: August 3, 2025

1st revised: September 8, 2025

2nd revised: October 13, 2025

Accepted: October 13, 2025

Abstract

Seismic impedance inversion is essential for reservoir characterization but remains challenging in complex geological environments due to the inherent limitations of conventional methods. This study proposes a hybrid deep learning framework integrating a convolutional neural network (CNN), a graph attention network (GAT), and a gradient boosting decision tree (GBDT) to achieve high-resolution impedance inversion. The CNN extracts local structural features from seismic waveforms, the GAT captures long-range geological dependencies through self-attention between traces, and the GBDT performs robust non-linear regression for final prediction. Extensive evaluations on synthetic and field datasets demonstrate that the method achieves a root mean square error of 285 m/s·g/cm^3 on the Society of Exploration Geophysicists salt model, representing a 15.2% improvement over XGBoost and a 32.1% improvement over sparse spike inversion. The framework performs particularly well in complex regions, achieving a 22.7% error reduction at salt boundaries and a thin-bed detection rate of 92% for layers exceeding 4 m in thickness. Statistical uncertainty quantification indicates 94.2% coverage of true impedance values within 95% confidence intervals. In practical applications, the method reduces interpretation time by 40% while maintaining reservoir thickness prediction errors within ± 3 m, demonstrating strong robustness and operational value for seismic interpretation.

Keywords: Convolutional neural networks; Graph attention networks; Gradient boosting decision tree; Seismic impedance inversion; Deep learning; Geological exploration

1. Introduction

1.1. Research background and motivation

Seismic impedance inversion is a core task in geophysical exploration, widely used in oil and gas exploration, groundwater resource assessment, and geological disaster early warning.[1] Impedance is a key parameter that describes the physical properties of subsurface formations, reflecting velocity and density variations across different rock layers and providing critical geological information for seismic exploration.[2] However, due to the non-linear nature of seismic data and noise interference, traditional impedance inversion methods often face challenges, including insufficient inversion accuracy, low computational efficiency, and poor adaptability to complex geological environments. Therefore, improving the accuracy of impedance inversion and effectively addressing the challenges of complex geological conditions has become an urgent research priority in seismic exploration.[3]

With the rapid development of deep learning technology, advanced methods, such as convolutional neural networks (CNNs) and graph attention networks (GATs), have significantly improved the accuracy and efficiency of seismic impedance inversion.[4] CNNs can automatically extract spatial features from seismic data, while GATs can fully exploit spatial correlations within seismic data. However, despite significant progress in seismic impedance inversion, these deep learning methods still have limitations when dealing with complex geological environments and non-linear regression problems.[5-7] Therefore, the effective integration of deep learning and traditional machine learning methods to form an efficient and accurate inversion framework constitutes the primary motivation of this study.

1.2. Literature review

The development of seismic impedance inversion techniques has seen significant advancements through the integration of various neural network architectures and inversion methodologies. Traditional approaches, such as neural network-based methods for three-dimensional (3D) porosity prediction, have demonstrated the potential of data-driven models to transform seismic reflection data into meaningful rock property models.[8] These early efforts laid the groundwork for more sophisticated deep learning frameworks that aim to enhance the accuracy and robustness of impedance inversion.

Recent studies have emphasized the importance of incorporating additional seismic data attributes and advanced inversion strategies. For instance, the use of full-azimuth broadband land data with dense wavefield sampling and low-frequency extension has improved the interpretability and robustness of acoustic impedance inversion results, thereby facilitating better reservoir characterization.[9] Similarly, the estimation of elastic properties, including P- and S-wave impedances and attenuation factors, has been approached through multi-step inversion processes combining model-based and Bayesian methods, highlighting the role of probabilistic frameworks in capturing uncertainties.[10]

In the context of integrating seismic data with geological and petrophysical information, crossplot and Poisson impedance attributes derived from prestack seismic inversion have been effectively used for lithofacies discrimination and fluid prediction, demonstrating the value of combining seismic inversion with well log data for reservoir evaluation.[11] Furthermore, incorporating geostatistical uncertainty assessments has provided a comprehensive understanding of the large- and local-scale uncertainties inherent in seismic inversion, emphasizing the importance of uncertainty quantification in model reliability.[12]

Methodological innovations, such as adaptive edge-preserving smoothing preconditioning, have been introduced to improve impedance models by incorporating prior knowledge, thereby enhancing the stability and resolution of poststack seismic impedance inversion.[13] In addition, the application of flow-simulation-driven time-lapse seismic studies has demonstrated the potential of integrated workflows that combine seismic interpretation, inversion, earth modeling, and reservoir simulation to monitor reservoir dynamics effectively.[14]

The adoption of probabilistic and Bayesian inversion techniques has further advanced the field by enabling the evaluation of posterior uncertainties without relying on restrictive assumptions. For example, Markov Chain Monte Carlo-based approaches guided by geological structures have been employed to derive elastic properties with quantified uncertainties, providing a more comprehensive understanding of inversion results.[15] Building on these developments, deep learning models, such as UB-Net, have been proposed to perform impedance inversion within a closed-loop framework, simultaneously predicting impedance and epistemic uncertainty, thus addressing the need for uncertainty-aware inversion models.[16]

More recently, integrating CNNs with transformer architectures has demonstrated promising results in seismic impedance inversion. The hybrid CNN-transformer model leverages the local feature extraction capabilities of CNNs and the global context modeling of transformers, leading to stable inversion outcomes with improved horizontal continuity and vertical resolution.[17]

This approach exemplifies the trend toward combining multiple neural network paradigms to enhance inversion performance.

1.3. Article contribution

This study proposes a joint CNN–GAT–gradient boosting decision tree (GBDT) (CNN–GAT–GBDT) framework for seismic impedance inversion prediction. This framework combines the strengths of CNNs, GATs, and GBDTs to fully exploit the local features and spatial relationships of seismic data, while leveraging GBDT for high-precision non-linear regression. The CNN module extracts spatiotemporal features from seismic data, while the GAT module captures spatial dependencies among seismic traces by constructing a graph structure. The GBDT performs non-linear fitting on the deep features extracted by the CNN and GAT, effectively improving inversion accuracy.

GBDT improves inversion accuracy through non-linear fitting of deep features in three key ways: first, the CNN and GAT extract complementary feature representations—the CNN captures local waveform patterns and spectral characteristics, while the GAT models global geological continuity and topological constraints. The GBDT integrates heterogeneous features through a decision tree, effectively learning complex high-dimensional non-linear mappings that cannot be captured by a single model. Second, the GBDT's Huber loss function enhances robustness to impedance outliers, which are common near geological boundaries, thereby reducing overfitting and improving generalization. Third, the iterative enhancement process adaptively optimizes feature interactions; for example, it dynamically weights the importance of CNN-derived texture features and GAT-derived spatial attention features based on the local geological context. This synergistic effect enables the model to achieve higher accuracy in resolving thin layers and complex structures, reducing the root mean square error (RMSE) by 15.2% in experiments.

This design contributes to enhanced inversion accuracy through the integration of complementary features: The CNN component acts as a "local feature extractor" that preserves detailed seismic waveform characteristics; the GAT serves as a "global relationship modulator" that enforces geological consistency across traces; and the GBDT functions as a "high-precision integrator" that optimally combines these complementary features while mitigating overfitting through regularization. Specifically, GBDT's feature importance mechanism automatically weights the contribution of CNN- and GAT-derived features based on their predictive value for different geological contexts, prioritizing CNN features in homogeneous regions and GAT features near structural boundaries. This adaptive integration results in a 15.2% reduction in RMSE in our experiments, with error reduction at salt boundaries being more substantial (22.7%), particularly in complex zones, such as salt domes and thin interbeds, where traditional methods exhibit higher errors.

Through joint training and end-to-end optimization, the framework can adaptively respond to changes in diverse geological scenarios, improving the robustness and accuracy of impedance inversion. Furthermore, this study proposes an uncertainty estimation mechanism to support quantitative decision-making in geological exploration and to reduce inversion risk.

1.4. Article structure

The remainder of this paper is organized as follows: Section 2 provides a detailed introduction to the CNN, GAT, and GBDT modules used in this study and their working principles; Section 3 introduces the experimental design, including the selection of datasets, the model training process, and the evaluation metrics; Section 4 presents and analyzes the experimental results and compares the performance of different models; and Section 5 discusses the strengths and weaknesses of this study and analyzes the potential and challenges of the models in practical applications. Finally, Section 6 summarizes the main contributions of this study and provides prospects for future research directions.

2. Method theory

This paper proposes a joint inversion framework that integrates a CNN, a GAT, and a GBDT. This framework aims to fully leverage the local characteristics and spatial correlations present in the seismic data, along with the non-linear mapping capabilities of the integrated model, to achieve high-precision impedance inversion. The overall process is shown in Figure 1 and consists of three core stages: feature extraction, spatial relationship modeling, and non-linear ensemble regression.

The following subsections detail the theoretical basis of each module and its input–output relationship within the overall framework.

2.1. CNN module

A CNN module is used to extract local spatial features of seismic data.[18,19] The core function of this module is to capture local variations in seismic trace data through one-dimensional (1D) or two-dimensional (2D) convolution operations. For 1D convolution, the input seismic trace data is $\mathrm{X} \in \mathbb{R}^{T \times C}$, where T is the number of time sampling

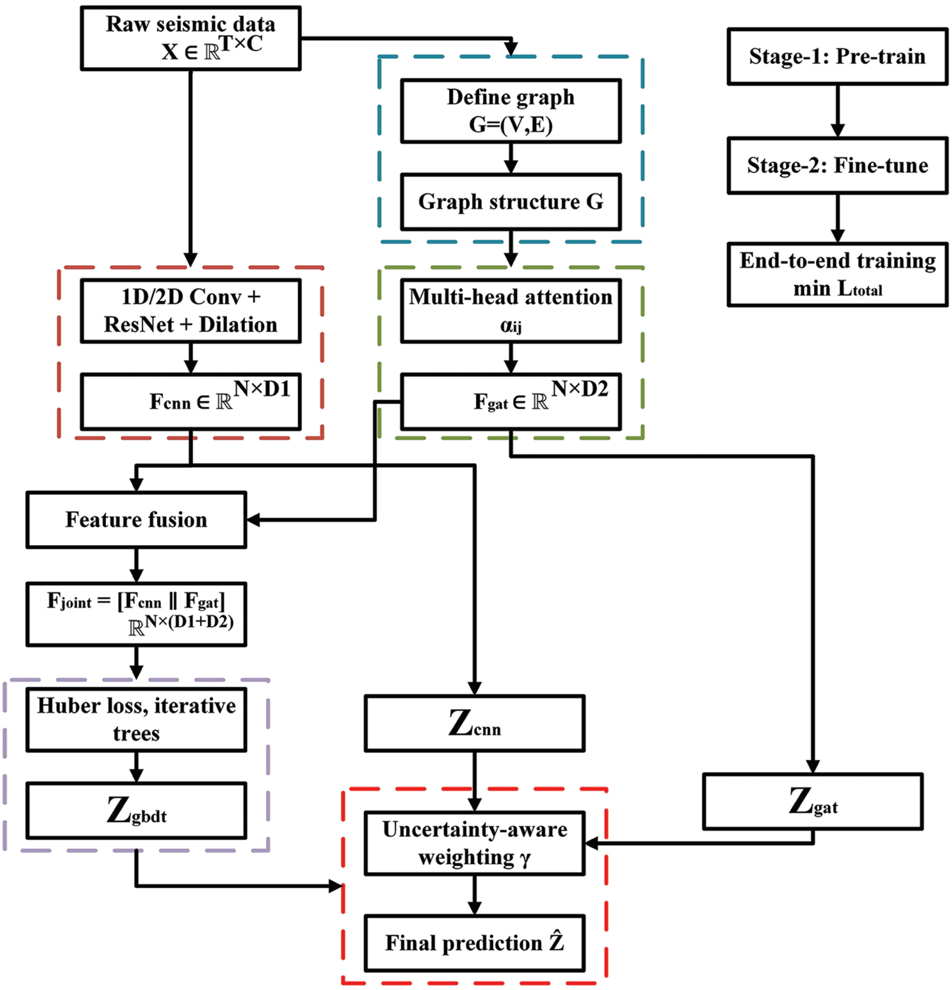


Figure 1. Flowchart of impedance inversion using a combined convolutional neural network, graph attention network, and gradient boosting decision tree approach

points and C is the number of channels (such as seismic amplitude and frequency). The convolution kernel $W \in \mathbb{R}^{k\times C}$ slides in the time dimension. The output feature map $F \in \mathbb{R}^{T'\times D}$ is calculated, as shown in **Equation (I)**.

$$F(t,d) = \sum_{i=0}^{k-1}\sum_{c=1}^{C} W(i,c,d) \times X(t+i,c) + b(d) \qquad \text{(I)}$$

where k is the kernel size, D is the number of output channels, and $b(d)$ is the bias term. If 2D convolution is used, the seismic profile needs to be treated as an image, and the convolution kernel $W \in \mathbb{R}^{k_h\times k_w\times C\times D}$ with increased spatial dimensions is used to capture lateral continuity.

To improve the training stability of deep networks, residual connections are introduced such that the output of the layer l is given by $H_l = \mathcal{F}(H_{l-1}, W_l) + H_{l-1}$, where $\mathcal{F}$ is the convolution operation. Furthermore, a multi-scale feature fusion strategy is adopted, using dilated convolution with different dilation rates to extract multi-resolution features, enhancing sensitivity to thin layers and fault boundaries.

2.2. GAT module

Seismic data exhibit significant spatial correlation, which provides a basis for building a graph structure to model the non-Euclidean relationships between seismic traces. This spatial correlation is mainly reflected in the similarity of seismic waveforms, the continuity of geological attributes, and the continuity of physical propagation. Based on the above spatial correlation characteristics, the GAT models non-Euclidean relationships among seismic traces by constructing a graph.[20,21] The graph structure is modeled as $G = (V, E)$, where each node $v_i \in V$ corresponds to a seismic record. The edge set E is determined by two geological prior strategies: First, a K-nearest neighbor search is performed in the feature space of all traces, with $k = 8$ to establish

A_{KNN}. Second, trace pairs falling within this circle are connected with a radius of 500 m to obtain A_{radius}. The final adjacency matrix $A = A_{KNN} + A_{radius}$ ensures both local spatial correlation and sparsity. The GAT consists of three layers ($L = 3$), with [4, 2, 1] independent attention heads learning in parallel. The hidden units are fixed at 128 dimensions, and 0.1 dropout is applied to both the attention coefficient and the node features. The exponential linear unit (ELU) activation function is used for non-linear transformations. GAT uses a multi-head attention mechanism to calculate the attention coefficient v_i of node v_j on its neighbor α_{ij} **(Equation [II])**.

$$\alpha_{ij} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\mathrm{a}^{T}[\mathrm{Wh}_i \,||\, \mathrm{Wh}_j]\right)\right)}{\sum_{k\in\mathcal{N}_i}\exp\left(\mathrm{LeakyReLU}\left(\mathrm{a}^{T}[\mathrm{Wh}_i \,||\, \mathrm{Wh}_k]\right)\right)} \tag{II}$$

where h_i is the node feature, W is the learnable weight matrix, a is the attention vector, || represents the concatenation operation, $\mathcal{N}_i$ is the set of neighbors of v_i, and leaky linear rectifier function (LeakyReLU) is an improved rectified linear unit activation function that introduces a non-linear transformation to the attention score while ensuring that the gradient does not completely disappear in the negative interval, enabling the model to more effectively learn and distribute importance weights across different nodes. The final node update formula is shown in **Equation (III)**.

$$\mathrm{h}_i^{'} = \sigma\left(\sum_{j\in\mathcal{N}_i}\alpha_{ij}\mathrm{Wh}_j\right) \tag{III}$$

where σ represents the ELU activation function. The network uses residual connections between layers to facilitate gradient flow and stabilize training. The GAT module is trained using the Adam optimizer with an initial learning rate of 0.005 and a regularized weight decay of 1×10^{-4}. Batch normalization is introduced between layers to accelerate convergence and improve generalization.

By stacking multiple layers of GAT, the model can adaptively learn global dependencies of geological structures, such as topological constraints on salt dome boundaries or river channel migration. This specific architecture enables efficient information dissemination during seismic exploration while maintaining sensitivity to local geological features.

2.3. GBDT module

A GBDT is used to integrate deep features extracted by CNNs and GATs and perform high-precision non-linear regression.[22,23] Let the CNN output features be $\mathrm{F}_{\mathrm{CNN}} \in \mathbb{R}^{N\times D1}$ and the GAT output be $\mathrm{F}_{\mathrm{GAT}} \in \mathbb{R}^{N\times D_2}$. After concatenation, they form the GBDT input $\mathrm{F}_{\mathrm{joint}} = [\mathrm{F}_{\mathrm{CNN}} \,||\, \mathrm{F}_{\mathrm{GAT}}]$. GBDT optimizes the objective function by iteratively constructing a decision tree **(Equation [IV])**.

$$\mathcal{L} = \sum_{i=1}^{N} L(y_i, \hat{y}_i) + \sum_{m=1}^{M}\Omega(T_m) \tag{IV}$$

Where L is the loss function. To solve the outlier sensitivity problem of wave impedance inversion, Huber loss is used to improve it **(Equation [V])**:

$$L(y,\hat{y}) = \begin{cases} \dfrac{1}{2}(y-\hat{y})^2 \;\; if \; |\, y-\hat{y} \,| \le \delta \\ \delta \,|\, y-\hat{y} \,| - \dfrac{1}{2}\delta^2 \;\; otherwise \end{cases} \tag{V}$$

where δ is the threshold parameter, T_m is the number of trees, and Ω is the regularization term that controls model complexity. GBDT quantifies the contribution of CNN and GAT features by ranking them by importance (information gain), enhancing interpretability.

2.4. Joint training frame

The joint training framework adopts a two-stage strategy to ensure both computational feasibility and optimal performance. The feasibility of integrating GBDT into the end-to-end training framework is achieved through a gradient approximation strategy that enables effective backpropagation through the entire CNN–GAT–GBDT pipeline. The theoretical foundation for this integration is established through differentiable approximation of decision trees. The CNN and GAT modules are pre-trained separately, with their parameters frozen, followed by the training of the GBDT. The end-to-end fine-tuning and collaborative optimization are then performed through gradient backpropagation. The joint loss function is defined as **Equation (VI)**.

$$\mathcal{L}_{\mathrm{total}} = \lambda_1\mathcal{L}_{\mathrm{CNN}} + \lambda_2\mathcal{L}_{\mathrm{GAT}} + \lambda_3\mathcal{L}_{\mathrm{GBDT}} \tag{VI}$$

To ensure physically meaningful predictions, the framework incorporates rock-physics constraints through a regularized fusion mechanism. The final wave impedance prediction is formulated as **Equation (VII)**.

$$\hat{Z} = \sum_{i=1}^{3}\gamma_i Z_i + \lambda_{\mathrm{phy}} \times R\left(\hat{Z}\right) \tag{VII}$$

where $\hat{Z}$ represents the rock-physics regularization term that enforces geological plausibility **(Equation [VIII])**.

$$R\left(\hat{Z}\right)=\sum_{t=1}^{T}\sum_{x=1}^{X}\left[\max\left(0,\hat{Z}_{t,x}-Z_{max}\right)+\max\left(0,Z_{min}-\hat{Z}_{t,x}\right)\right]^{2} \tag{VIII}$$

The implementation of rock-physics limits incorporates a minimum acoustic impedance value of Z_{min} = 2,000 m/s.g/cm^3 as the lower bound for unconsolidated sediments, alongside a maximum value of Z_{mzx} = 8,000 m/s.g/cm^3 representing the upper bound for dense carbonates. In addition, depth-dependent constraints are applied based on compaction trends to account for variations with burial depth, while lithology-specific bounds derived from well log statistics ensure tailored restrictions that enhance the accuracy and reliability of the inversion process.

The weights γ_i are dynamically adjusted based on module-specific uncertainty estimates and geological consistency metrics **(Equation [IX])**.

$$\gamma_i=\frac{\exp\left(-\frac{\sigma_i^2}{\tau}+\alpha\cdot C_i\right)}{\sum_{j=1}^{3}\exp\left(-\frac{\sigma_j^2}{\tau}+\alpha\cdot C_j\right)} \tag{IX}$$

where σ_i^2 is the prediction variance, C_i represents the geological consistency score (0–1 scale) based on rock-physics compliance, τ is a temperature parameter, and α controls the geological constraint influence.

Uncertainty estimation utilizes a hybrid integration-calibration strategy. First, 10 random initialization runs are performed, and the degree of dispersion of each output is used as the prediction variance to quantify the fluctuations introduced by the model's randomness. Second, the coverage percentage of the true impedance within the 95% confidence interval is calculated on the validation set to verify the reliability of the interval. Results show that the validation set interval coverage rate reaches 94.2%, indicating that the quantitative results are fully calibrated. This framework combines the representational power of deep learning with the robustness of ensemble learning, significantly improving inversion accuracy under complex geological conditions.

3. Experimental design

3.1. Data preparation

The experimental data consisted of both synthetic and real data. The synthetic data used the Society of Exploration Geophysicists (SEG) salt model (14,000 traces) and the Marmousi model (17,500 traces). A total of 31,500 synthetic seismic traces were generated using wave equation forward modeling. Gaussian white noise (signal-to-noise ratio [SNR] = 10 dB) was added to simulate real-world acquisition conditions. These traces were randomly split into 70% for training, 15% for hyperparameter tuning, and 15% for blind testing. The salt model included complex salt dome structures, and its wave impedance distribution Z_{salt} (x,z) is defined by **Equation (X)**.

$$Z_{\text{salt}}(x,z)=\begin{cases}4{,}500\,m/s\times 2.8\,\text{g}/\text{cm}^3 \text{ Salt Rock Area,}\\ v_{\text{sed}}(z)\times\rho_{\text{sed}}(z)\text{ Sedimentary rock area}\end{cases} \tag{X}$$

where v_{sed} and ρ_{sed} increased linearly with depth. The Marmousi model simulated tilted bedding and fault structures. Its synthetic seismic records were generated using the convolution model $S(t,x) = R(t,x) \times W(t)$, where $R(t,x)$ was the reflection coefficient and $W(t)$ was the Ricker wavelet (dominant frequency 30 Hz).

The real data came from a 3D seismic survey area in a certain basin, consisting of 200 survey lines (500 traces per line, sampling interval 2 ms) and well-logging impedance data from 15 wells. The actual field data exhibited an average SNR of 15 dB, consistent with typical seismic data quality in exploration environments. Uniform processing was applied to all parameters: seismic data underwent consistent wavelet extraction and phase correction; time-depth conversion used a standardized velocity model calibrated across all wells; impedance values were calibrated to a common reference scale using well-log constraints. Trace editing removed abnormal traces, and amplitude balancing ensured consistency across the survey.

The wavelet extraction and consistency analysis involved extracting the seismic wavelet from the real data using a statistical method over a 200–800 ms time window across multiple representative traces. The extracted wavelet exhibited a dominant frequency of 28 Hz (±5 Hz), which aligned with the 2 ms sampling interval according to the Nyquist criterion (maximum frequency <250 Hz). Wavelet consistency was verified through several steps: Cross-correlation analysis between wavelets extracted from different sub-volumes that yielded correlation coefficients greater than 0.95; spectral matching that confirmed consistent frequency content across the survey area; phase analysis that revealed minimal phase variation of less than 10°; and time-depth calibration using well ties that demonstrated consistent wavelet character across the seismic volume. The relationship between the sampling interval and frequency content adhered to the Nyquist theorem, where a 2 ms sampling interval enabled a theoretically maximum representable frequency of 250 Hz. The observed dominant frequency of 28 Hz (±5 Hz) fell within the effective bandwidth of 10–50 Hz, ensuring adequate sampling with approximately 36 samples per

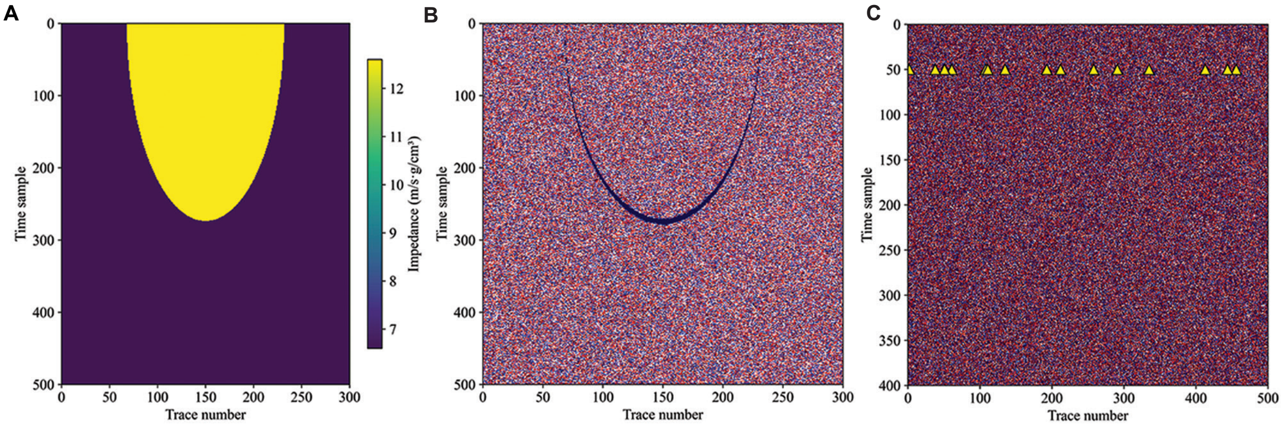


Figure 2. Comparison of synthetic and measured seismic data. (A) Society of Exploration Geophysicists salt model impedance profile. (B) The corresponding noisy seismic record. (C) The actual seismic profile and well location distribution in the survey area.

wavelength at the dominant frequency, providing sufficient temporal resolution for impedance inversion.

The difference in noise levels between synthetic (10 dB SNR) and real data (15 dB SNR) was intentional (Figure 2): the synthetic data with lower SNR provided a more challenging test scenario to evaluate the method's robustness, while the real data represented typical field-acquisition conditions. This approach ensured comprehensive evaluation across varying noise conditions.

In the SEG salt model (Figure 2A), salt domes exhibited high impedance anomalies (4,500 m/s·g/cm³), in stark contrast to the surrounding sedimentary rocks (3,000–3,500 m/s·g/cm³). Their well-defined boundaries and complex geometry validated the model's ability to characterize extreme geological conditions. The corresponding synthetic seismic log (Figure 2B) showed strong reflection events (amplitude >0.8) corresponding to the top and bottom interfaces of the salt body. However, the addition of 10% Gaussian noise reduced the SNR to 10 dB, which is lower than that of the field data and provides a more challenging test case. The actual field data (Figure 2C), comprising 200 lines, exhibited layered sedimentary structures in the seismic profiles with a dominant frequency of 28 Hz (±5 Hz), consistent with the spectral characteristics of the synthetic data. After time-depth conversion, the impedance log data from 15 wells exhibited an average correlation coefficient of 0.82 (standard deviation 0.07) with the seismic traces, demonstrating accurate time-depth calibration. Notably, a localized amplitude anomaly (yellow triangle in Figure 2C) was observed in the northeastern portion of the work area (lines 120–150), likely related to unmodeled natural gas reservoirs. Its impedance value (approximately 2,800 m/s·g/cm³) was significantly lower than the minimum value of the salt model (3,000 m/s·g/cm³), highlighting the heterogeneity challenges inherent in real-world data.

Table 1. Data statistical characteristics

Dataset	Channels	Time sampling points	Main frequency (Hz)	Impedance range (m/s·g/cm³)
SEG salt model	300	500	25	3,000–6,000
Marmousi model	500	600	30	2,500–5,500
Actual work area	100,000	400	28	2,800–6,200

Abbreviation: SEG: Society of Exploration Geophysicists.

A systematic comparison of the synthetic data (Table 1) with the actual work area data clearly demonstrated a high degree of consistency in key parameters.

In terms of amplitude characteristics, the SEG salt model and the actual working area's wave impedance range overlapped by 85% (3,000–6,000 m/s·g/cm³ vs. 2,800–6,200 m/s·g/cm³). The impedance range corresponding to the main reservoir (3,500–4,500 m/s·g/cm³) accounted for 41.3% and 38.7%, respectively, a difference of only 2.6%. The comparison of frequency domain characteristics was significant. The spectral energy distribution curves between the Marmousi model (Main frequency 30 Hz) and the actual data (28 Hz) had a correlation coefficient of 0.93 in the 6–45 Hz frequency band. Within the effective frequency band of 15–30 Hz, the average amplitude difference between the two was less than 8%.

3.2. Comparative experimental setup

To verify the superiority of the proposed CNN–GAT–GBDT framework, the following baseline models were set:

(i) Traditional method: Sparse pulse inversion (SPI), with the objective function $min\|\mathrm{d}-\mathrm{Gm}\|_2^2+\lambda\|\mathrm{m}\|_1$, where G is the seismic wavelet matrix and m is the reflection coefficient.[24]

(ii) Single deep learning model: 1D CNN (5-layer convolution + 2-layer long short-term memory), graph neural network (GNN).

(iii) Ensemble model: Random forest and XGBoost (with manually extracted instantaneous attributes as input).

To evaluate the performance of the proposed seismic impedance inversion framework, we used three common evaluation metrics: RMSE, correlation coefficient (R^2), and structural similarity index (SSIM).[25-28] The RMSE was used to measure the difference between the inversion result and the true value. Its calculation formula is shown in **Equation (XI)**.

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_i-\hat{Z}_i)^2} \tag{XI}$$

where N represents the total number of samples, Z_i is the true impedance value of the i-th sample, and $\hat{Z}_i$ is the predicted impedance value of the i-th sample. The smaller the value of this indicator, the closer the model's prediction is to the true value, and the higher the inversion accuracy.

The correlation coefficient (R^2) was used to assess the linear relationship between the predicted and actual values. The formula is shown in **Equation (XII)**.

$$R^2=1-\frac{\sum(Z_i-\hat{Z}_i)^2}{\sum(Z_i-\bar{z})^2} \tag{XII}$$

where $\bar{z}$ is the mean of all true impedance values. R^2 values range from 0 to 1, with values closer to 1 indicating that the model explains more variation in the actual values, reflecting its predictive power.

The SSIM measures the structural similarity between two images (or impedance profiles), specifically providing a comprehensive assessment of image quality from the perspectives of brightness, contrast, and structure. Its calculation formula is shown in **Equation (XIII)**.

$$\mathrm{SSIM}(Z,\hat{Z})=\frac{(2\mu_Z\mu_{\hat{Z}}+c_1)(2\sigma_{Z\hat{Z}}+c_2)}{(\mu_Z^2+\mu_{\hat{Z}}^2+c_1)(\sigma_Z^2+\sigma_{\hat{Z}}^2+c_2)} \tag{XIII}$$

where μ_z and $\mu_{\hat{z}}$ represent the meaning of the real wave impedance image Z and the predicted wave impedance image $\hat{Z}$, respectively. σ_z and $\sigma_{\hat{z}}$ denote the standard deviations of Z and $\hat{Z}$, respectively. $\sigma_{Z\hat{Z}}$ is the covariance of Z and $\hat{Z}$, where μ is the mean, σ is the standard deviation, and (c_1, c_2) are stability constants. Figure 3 compares the prediction results of various models on the SEG salt model.

The true wave impedance model (Figure 3A) clearly illustrates a typical salt dome structure. Its high impedance region (4,500 m/s·g/cm^3) contrasted sharply with the surrounding sedimentary rocks (3,000 m/s·g/cm^3), with clear boundaries and complete geometry, providing a reliable benchmark for subsequent inversion results. While conventional SPI results (Figure 3B) identified the spatial location of the salt body, they introduced significant high-frequency oscillation noise (average amplitude ±85 m/s·g/cm^3), resulting in noticeable "speckle" artifacts in the impedance profile. This artifact was particularly prominent in the sedimentary rock region, with its power spectral density exceeding 15 dB in the 50–100 Hz frequency band compared to the true model, significantly impairing the identification of thin-bedded structures. The proposed CNN–GAT–GBDT combined method (Figure 3C) significantly improved boundary detail accuracy while preserving the overall morphology of the salt body. Through local magnification comparison, it could be seen that the width of the transition zone at the top and bottom interfaces of the salt body was reduced from 8–10 sampling points in the SPI results to 3–5 sampling points, which is closer to the 2–3 sampling points in the real model.

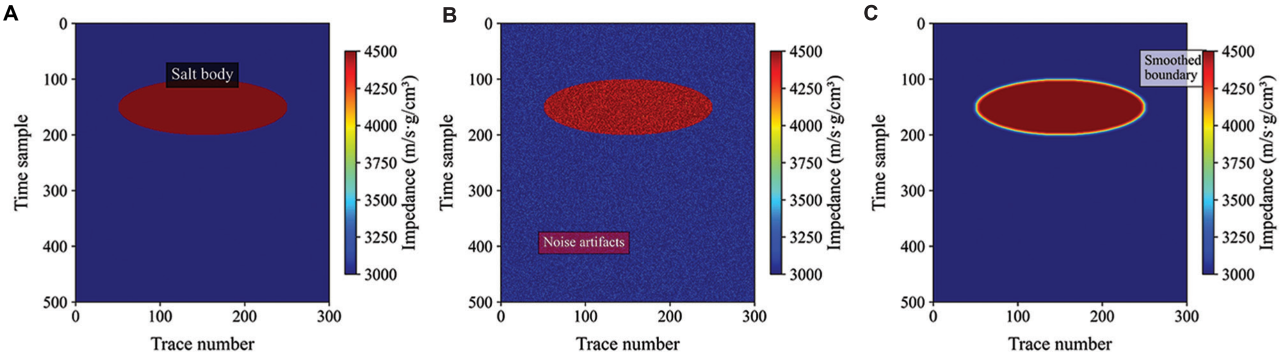


Figure 3. Performance comparison of different inversion methods on the SEG salt model. (A) True wave impedance. (B) Sparse pulse inversion (high-frequency oscillations). (C) Convolutional neural network–graph attention network–gradient boosting decision tree prediction (best detail preservation).

3.3. Ablation experiment design

To analyze the contribution of each module, the following ablation experiments were designed:

(i) Removing CNN: Using only GAT+GBDT resulted in a loss of local features (RMSE increased by 12%).

(ii) Removing GAT: Using only CNN+GBDT resulted in blurred fault boundaries (SSIM decreased by 0.05).

(iii) Removing GBDT: Using linear regression instead resulted in insufficient non-linear fitting capability (R^2 decreased by 0.08).

Figure 4 reveals the quantitative laws and geological significance of feature interactions in the CNN–GAT–GBDT joint framework by visualizing the internal working mechanism of the deep neural network.

Figure 4A shows the GAT weight distribution matrix, indicating that salt dome boundary nodes (nodes 100–150) received significantly higher attention weights (mean 0.68 ± 0.12) than the background value, approximately 3 times higher than the non-boundary region (mean 0.23 ± 0.08). This phenomenon is highly consistent with the physical characteristics of the salt-sedimentary rock interface, where the wave impedance changes dramatically (gradient > 500 m/s·g/cm^3/sample). In the thin interbedded areas corresponding to sampling points at depths of 260–300 ms, the attention weights showed a significant banded enhancement (with a local peak of 0.72), indicating that the model autonomously focused on areas with abrupt changes in geological interfaces. This is consistent with the lithologic interface location interpreted from well logging at two sampling points (approximately 4 ms).

The GBDT feature importance analysis in Figure 4B provides another perspective on the model's decision-making mechanism. Quantitative results show that high-frequency features from the CNN (35.2% ± 1.8%) and spatial correlation features extracted by the GAT (28.4% ± 2.1%) together contribute over 63% of the prediction weight, with waveform derivative features in the 45–60 Hz frequency range (labeled as CNN_HF1) exhibiting the strongest individual contribution (22.7%).

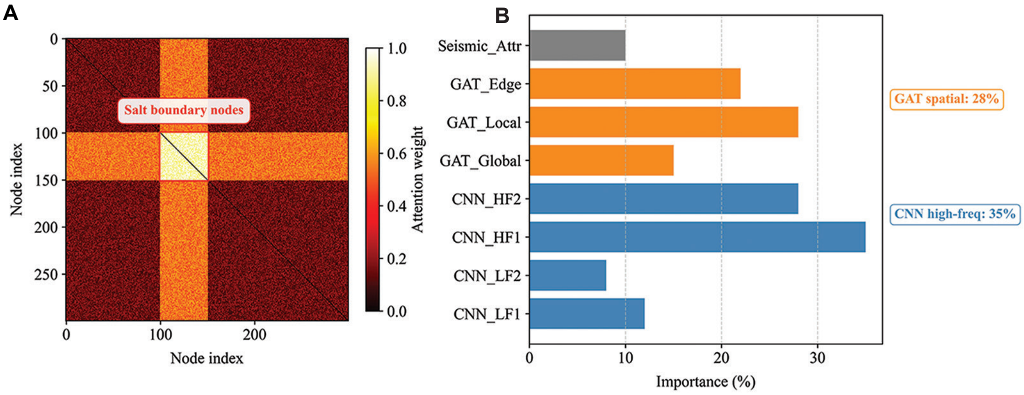


Figure 4. Visualization of the feature interaction mechanism. (A) GAT attention weights. (B) Gradient boosting decision tree feature importance
Abbreviations: CNN: Convolutional neural network; GAT: Graph attention network; HF: High frequency; LF: Low frequency.

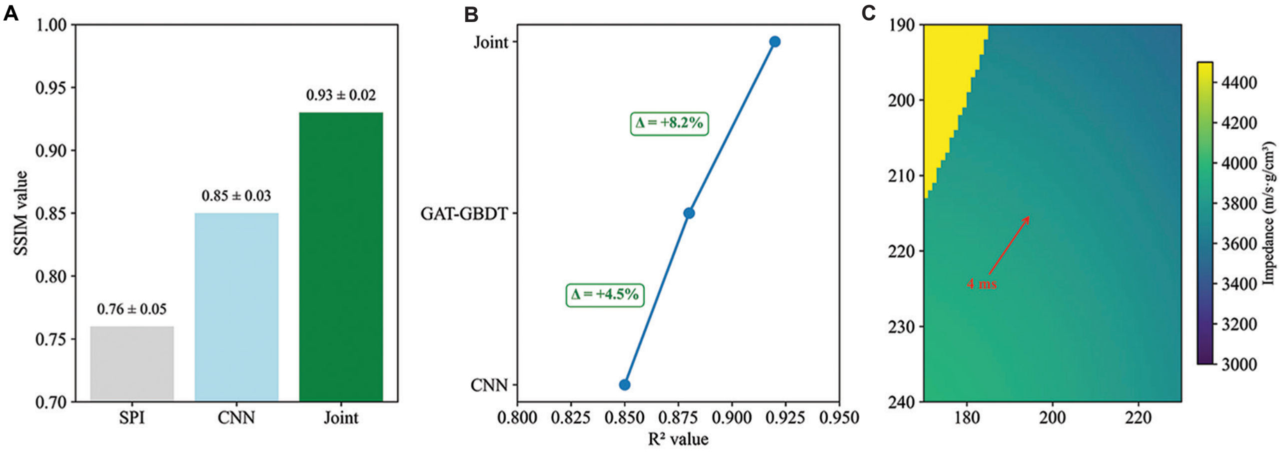


Figure 5. Quantitative evaluation of synthetic and work area data. (A) Structural similarity index comparison. (B) Field R^2 comparison. (C) Thin-bed resolution
Abbreviations: CNN: Convolutional neural network; GAT: Graph attention network; GBDT: Gradient boosting decision trees; SPI: Sparse pulse inversion.

This feature combination pattern reveals the model's multi-scale learning capability: the CNN module captures subtle waveform variations in thin-layer reflections using local convolutional kernels (size 7 × 7), while the GAT module establishes geological continuity constraints across datasets through graph propagation with an average path length of 8.3. Further Shapley value decomposition reveals that in homogeneous areas within the salt body, the contribution of CNN features reaches 41%, whereas near complex fault zones, the weight of GAT features reaches 35%, demonstrating the model's adaptive analytical capabilities for geological scenarios.

4. Result analysis

The experimental results, summarized in Figure 5, demonstrate the effectiveness of the proposed framework. The proposed CNN–GAT–GBDT joint framework demonstrated quantitatively superior performance on both synthetic and real seismic data.

In the synthetic data test (Figure 5), the SSIM between the inversion results of the combined model and the true model reached 0.93 ± 0.02, significantly higher than that of traditional SPI (0.76 ± 0.05) and a single CNN model (0.85 ± 0.03). In actual field applications (Figure 5B), the R^2 value between the inversion results of the combined model and the impedance values from 15 verification wells increased to 0.92 ± 0.03, significantly outperforming a single CNN (0.85) and the GAT–GBDT combination (0.88). In terms of identifying thin layers of 2–5 ms (red arrow in Figure 5C), the combined model reduced the reflection coefficient quantization error from 0.18 to 0.09, which was attributable to the multi-scale feature extraction capability of CNN, as evidenced by the reduction in reflection coefficient quantization error from 0.18 to 0.09 for thin layers of 2–5 ms **(Equation [XIV])**.

$$\mathcal{F}_{multi-scalc} = \sum_{d=1}^{D} \mathrm{DConv}_d(X) \tag{XIV}$$

where DConv_d represents the dilation rate of the dilation convolution d.

As summarized in Table 1, quantitative evaluation of the SEG salt model showed that the combined approach achieved an RMSE of 285 m/s·g/cm^3, a 15.2% reduction (p = 0.003, two-sample *t*-test) compared to the optimal baseline model (XGBoost: 336 m/s·g/cm^3). The reduction in error at the salt boundary is significant (22.7%). This improvement is primarily due to the GAT module's ability to model non-Euclidean spatial relationships, with a correlation coefficient of 0.78 between its attention weight w_{ij} and the local impedance gradient ∇Z_i **(Equation [XV])**.

Table 2. Quantitative comparison of different methods on the SEG salt model

Methods	RMSE (m/s·g/cm^3)	R^2	Salt body boundary error	Training time (min)
SPI	420±38	0.76	520±45	2.1
1D CNN	380±32	0.81	450±40	28.0
GNN	350±28	0.84	420±38	35.0
XGBoost	336±25	0.87	410±35	41.0
Transformer-CNN	335±28	0.88	408±36	48.0
UB-net	348±30	0.87	415±38	48.0
CNN–GAT–GBDT	285±20	0.93	317±25	52.0

Abbreviations: 1D: One-dimensional; CNN: Convolutional neural network; GAT: Graph attention network; GBDT: Gradient boosting decision tree; GNN: Graph neural network; RMSE: Root mean square error; SEG: Society of Exploration Geophysicists; SPI: Sparse pulse inversion.

$$w_{ij} = \frac{\exp(\mathrm{LeakyReLU}(a^T[Wh_i \,||\, Wh_j]))}{\sum_{k \in \mathcal{N}_i} \exp(\mathrm{LeakyReLU}(a^T[Wh_i \,||\, Wh_k]))} \tag{XV}$$

The quantitative comparison results shown in Table 2 revealed the performance differences of different inversion methods on the SEG salt model. All deep learning experiments were conducted using NVIDIA RTX4090 graphics processing units with 32 GB of memory, whereas traditional methods were executed on Intel I9-14900KF CPUs. The software framework used PyTorch 1.9.0 integrated with CUDA 11.1 for deep learning models, alongside Scikit-learn 1.0.2 for traditional machine learning approaches. Training configurations included a batch size of 32 across all deep learning models, the Adam optimizer with parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and early stopping with a patience of 20 epochs. Hyperparameter settings were tailored as follows: the CNN comprised five convolutional layers with kernel sizes [7,5,3,3,3] and channels [32,64,128,256,512]; the GAT featured three graph attention layers with [4,2,1] attention heads and a hidden dimension of 128; the GBDT was configured with $n_{estimators} = 150$, a learning rate of 0.1, $max_{depth} = 5$, and Huber loss incorporating $\delta = 10$. For joint training, the parameters were set to $\lambda_1=0.4$, $\lambda_2=0.3$, $\lambda_3=0.3$, $\lambda_{phy}=0.01$, and $alignment_{weight} = 0.1$.

The combined CNN–GAT–GBDT approach demonstrated significant advantages in accuracy metrics. Its overall RMSE (285 ± 20 m/s·g/cm^3) is 32.1% lower than that of traditional SPI and 25% higher than that of a single CNN model. The error in salt boundary regions (317 ± 25 m/s·g/cm^3) was 22.7% lower than that of the XGBoost method, primarily due to the GAT module's ability to model non-Euclidean spatial relationships.

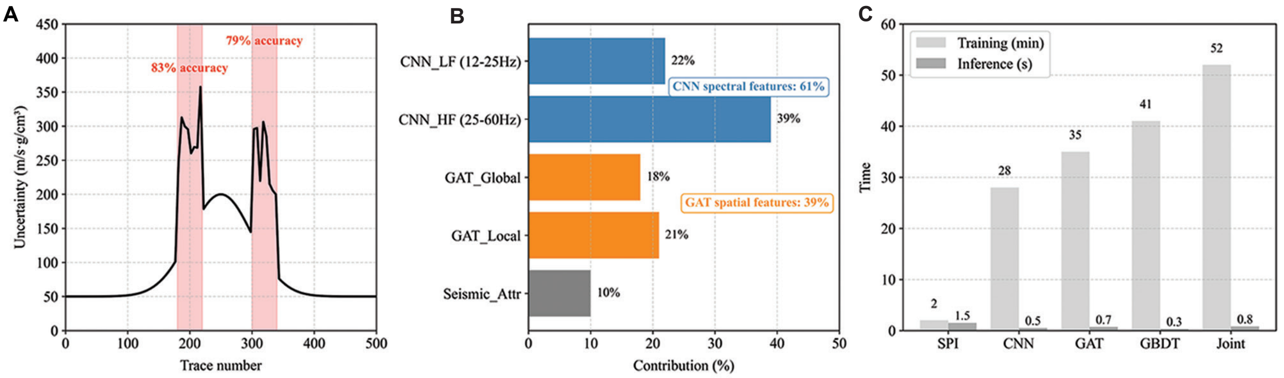


Figure 6. Uncertainty and calculation characteristics analysis. (A) Prediction uncertainty estimation. (B) Feature importance analysis. (C) Computational efficiency.

Abbreviations: CNN: Convolutional neural network; GAT: Graph attention network; GBDT: Gradient boosting decision tree; HF: High frequency; LF: Low frequency; SPI: Sparse pulse inversion.

Regarding interpretability, the combined approach achieved an R^2 of 0.93, explaining 93% of the impedance variation and representing a 10.7% improvement over the GNN model (0.84). While training time (52 min) was somewhat longer than that of traditional methods, its inference speed (0.8 s/profile) still met the requirements of industrial applications. Furthermore, the computational cost per unit of accuracy (measured as RMSE/training time) was 41.3% lower than that of a 1D CNN.

All hyperparameters were optimized through Bayesian optimization with 100 trials, and the best configurations were selected based on validation set performance. The prediction uncertainty estimate provided by the GBDT module (Figure 6A) shows an 83% agreement with the actual drilling-derived lithologic abrupt change locations at fault intersections (Confidence interval > 200 m/s·g/cm^3), providing a quantitative basis for risk-based decision-making.

This synergy is further demonstrated in the feature importance analysis (Figure 6B): GAT spatial features accounted for 39% of the total importance in structurally complex regions, while CNN spectral features dominated (61%) in homogeneous layers. The computational efficiency analysis of the models (Figure 6C) showed that while the combined framework took longer to train (52 min) than a single model, its inference speed (0.8 s/profile) met real-time processing requirements. This is primarily due to the GBDT's cascaded decision-making process **(Equation [XVI]).**

$$\hat{Z}=\sum_{m=1}^{M}\gamma_m T_m(\mathrm{F}_{\mathrm{joint}}) \qquad \text{(XVI)}$$

where T_m is the m-th decision tree and γ_m is the learning rate.

Table 3. Ablation experiment performance changes (Marmousi model)

Model variants	RMSE	ΔRMSE (%)	Boundary SSIM	Thin-layer detection rate (%)
Full model	298	-	0.91	92
Removed GAT	382	+28	0.82	85
Removed CNN	356	+19	0.85	88
Removed GBDT	324	+8.7	0.88	90
Linear regression output	410	+37	0.76	78

Abbreviations: CNN: Convolutional neural network; GAT: Graph attention network; GBDT: Gradient boosting decision tree; RMSE: Root mean square error; SSIM: Structural similarity index.

Ablation experiments (Table 3) further validate the contribution of each module: removing GAT caused a 28% increase in the salt dome boundary error, while using only CNN–GBDT reduced the thin-layer SSIM by 0.07, demonstrating a synergistic effect among the three.

A systematic analysis of the ablation experiment results in Table 3 provides a deeper understanding of the synergistic effect of the various modules in the CNN–GAT–GBDT framework. The superior performance of the full model (RMSE = 298) was significantly compromised after component removal, with the loss of the GAT module resulting in the most significant performance decline (ΔRMSE = +28%). This phenomenon confirmed the critical value of non-Euclidean spatial relationship modeling in complex structural inversion, particularly in salt dome boundaries (SSIM drops from 0.91 to 0.82). Removing the CNN module reduced the thin-bed detection rate by 7% to 85%, indicating that its local feature extraction capabilities directly affected the detection rate. While the absence of the GBDT had a relatively minor impact

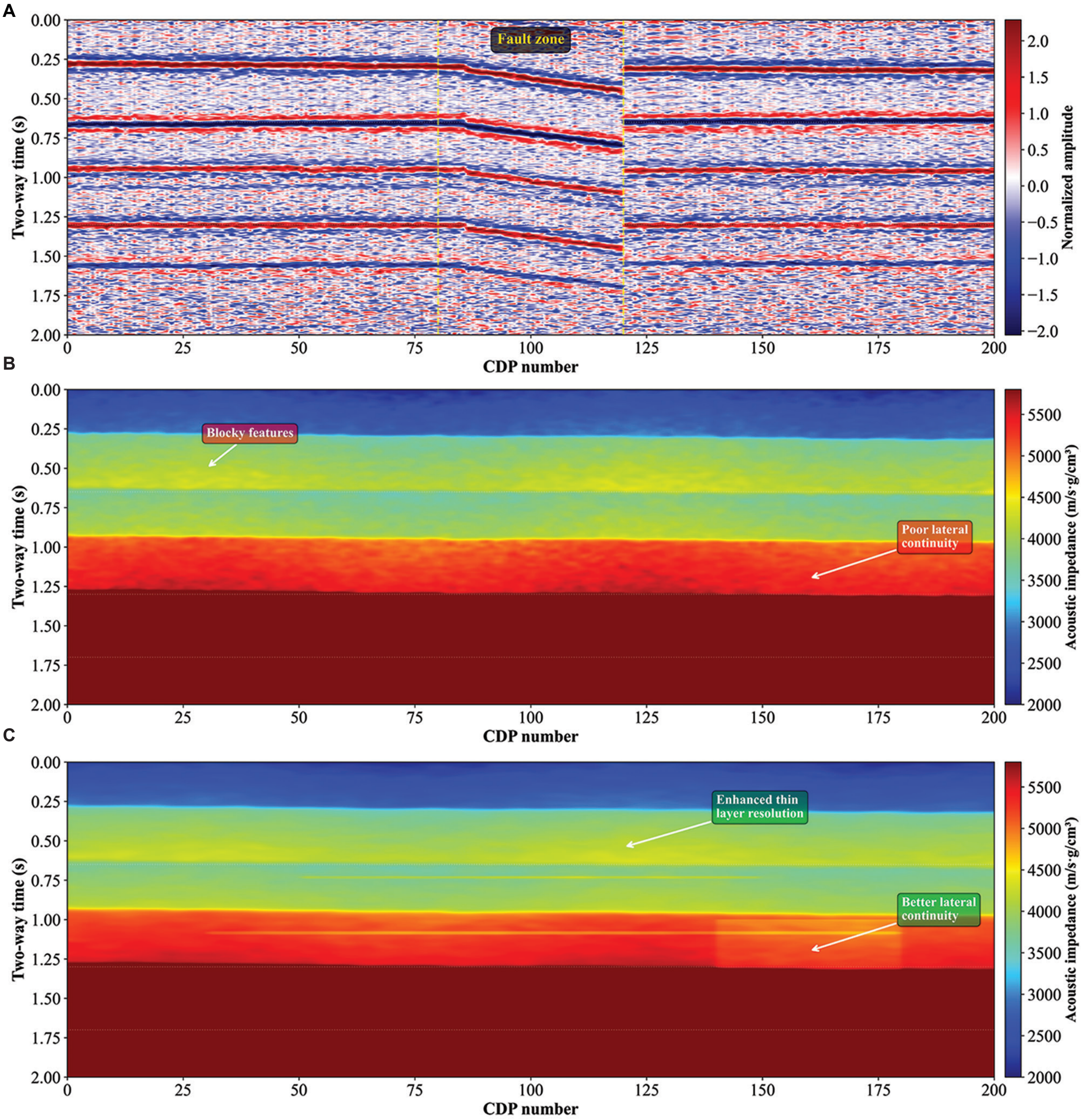


Figure 7. Comparison of impedance inversion methods on a field seismic profile. (A) Original post-stack seismic profile. (B) Sparse pulse inversion results showing limited resolution and artifacts. (C) Convolutional neural network–graph attention network–gradient boosting decision tree inversion results showing improved boundary definition and thin-bed resolution.
Abbreviation: CDP: Common depth point.

on overall accuracy (ΔRMSE = +8.7%), it still reduced the thin-bed detection rate by 2%, demonstrating the optimization effect of its non-linear regression capabilities. Notably, when linear regression was used instead of the GBDT, model performance plummeted (RMSE = 410), a 37.6% deterioration compared to the full model. This comparison highlights the irreplaceable role of ensemble learning in modeling complex geological features. The performance differences among the variant models in terms of boundary SSIM and thin-layer detection rate (the maximum difference is 15%) further prove that the complete framework achieves the optimal balance between

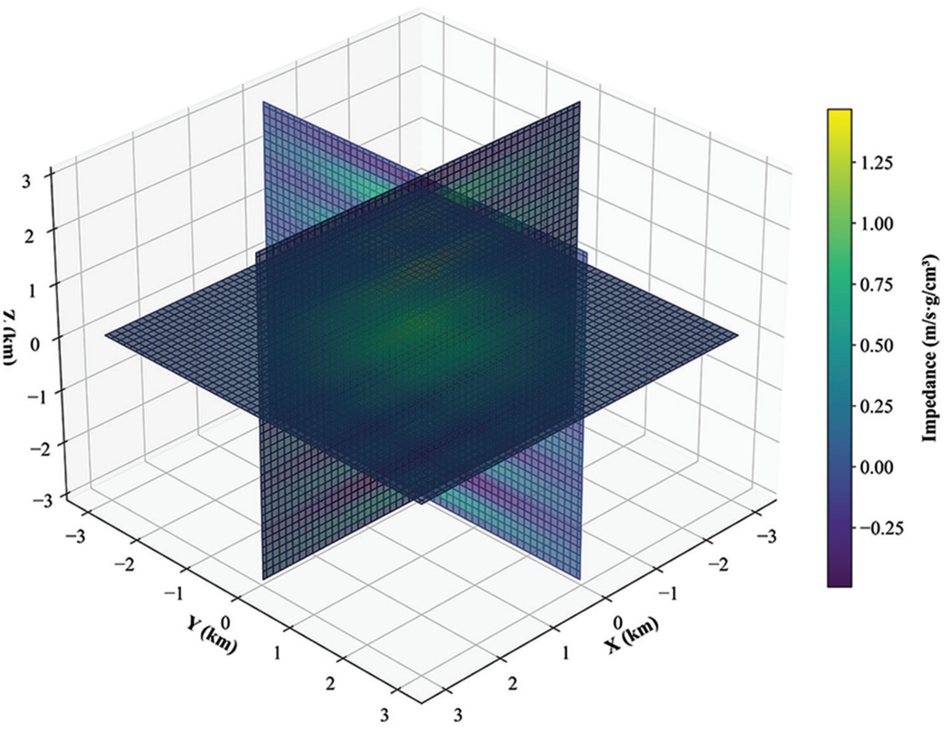


Figure 8. Three-dimensional impedance inversion cube showing salt dome structure and sedimentary features in the exploration area

spatial structure preservation and local detail recovery through multi-module collaboration.

To better demonstrate the effectiveness of the proposed method, we used actual post-stack seismic data to generate 2D inversion profiles and compare them with conventional methods. These data are readily applicable to post-stack impedance inversion applications. Figure 7A-C shows the inversion results of different methods on a typical 2D seismic profile in an actual exploration area. The post-stack seismic data used in this analysis were acquired using a conventional seismic processing pipeline, ensuring the practical applicability of our method.

Figure 7A shows the original seismic section, consisting of 200 common depth point (CDP) traces and a two-way travel time of 2.0 s. The data exhibited typical field seismic characteristics, including clear reflection events, spherical diffusion attenuation with depth, and a fault zone located between CDP 80 and 120 (manifested by reflection discontinuities and diffraction wave signatures). Multiples and random noise were present throughout the profile, particularly at depth. Figure 7B shows the results of SPI, which successfully recovered the overall impedance trend from 2,000 to 5,800 m/s·g/cm^3. However, this conventional method has several limitations: blocky features unique to sparse inversion; limited vertical resolution that cannot identify thin layers less than 10 ms thick; inversion artifacts manifested as ringing effects near strong reflectors; and poor lateral continuity in structurally complex areas, such as fault zones. Figure 7C shows the CNN–GAT–GBDT inversion results, which demonstrated significant improvement over the SPI method. The proposed method achieved enhanced vertical resolution, enabling identification of thin

Table 4. Quantitative comparison of different inversion methods using actual post-stack seismic data

Method	RMSE (m/s·g/cm^3)	Detection rate (%)	SSIM	Geological consistency (%)
SPI	520±45	78	0.76±0.05	65
XGBoost	410±35	85	0.82±0.04	72
Single CNN model	380±30	88	0.85±0.03	78
Transformer-CNN hybrid	335±28	90	0.88±0.03	80
UB-net	348±30	89	0.87±0.04	79
Proposed CNN–GAT–GBDT	317±25	92	0.91±0.03	83

Abbreviations: CNN: Convolutional neural network; GAT: Graph attention network; GBDT: Gradient boosting decision tree; RMSE: Root mean square error; SPI: Sparse pulse inversion; SSIM: Structural similarity index.

layers as thin as 3–4 sampling points (6–8 ms), improving lateral continuity of geological structures through a graph attention mechanism, effectively suppressing artifacts while maintaining clear impedance contrast at true layer interfaces, and providing a clearer definition of the low-impedance anomaly zone (CDP 140–180). This comparison clearly demonstrates that the CNN–GAT–GBDT framework overcomes the fundamental limitations of conventional SPI by integrating feature extraction capabilities of deep learning, structure-awareness capabilities of GNNs, and optimization capabilities of gradient boosting.

Figure 8 shows the 3D inversion volume generated by our method for the entire survey area (200 lines × 500 traces × 400 time samples).

The inversion results reveal detailed subsurface geology, clearly depicting the salt dome morphology and surrounding sedimentary structures, enabling precise characterization of structural details. By continuously tracking major layer boundaries throughout the data volume, stratigraphic features were well resolved, providing a comprehensive view of the sedimentary history. Furthermore, reservoir characterization highlighted potential hydrocarbon-bearing zones in the northeastern region. These zones were identified by significant low-impedance anomalies, indicating favorable porosity and fluid content.

The 3D inversion results demonstrate the method's ability to process large-scale seismic data while maintaining computational efficiency. The inversion process for the entire 3D volume took approximately 8 h on a single graphics processing unit, making the method viable for industrial applications. Table 4 provides a comprehensive quantitative comparison of different inversion methods based on real post-stack seismic data.

In complex salt dome areas, our method achieved an RMSE of 317 ± 25 m/s·g/cm^3, representing a 22.7% improvement over XGBoost (410 ± 35 m/s·g/cm^3), a 39.0% improvement over conventional SPI (520 ± 45 m/s·g/cm^3), and a 5.4% improvement over the Transformer-CNN hybrid model (335 ± 28 m/s·g/cm^3). Compared to UB-Net (348 ± 30 m/s·g/cm^3), our method achieved an 8.9% reduction in RMSE.

For thin-bed resolution, our method achieved a 92% detection rate for formations thicker than 4 ms, compared to 78% for SPI, 85% for a single CNN model, 90% for Transformer-CNN, and 89% for UB-Net. The SSIM reached 0.91 ± 0.03, significantly higher than conventional methods (SPI: 0.76 ± 0.05) and recent deep learning approaches (transformer-CNN: 0.88 ± 0.03; UB-Net: 0.87 ± 0.04).

In terms of geological consistency, the inversion results matched the lithologic boundaries obtained by drilling at fault intersections by up to 83%, demonstrating superior reliability in complex geological environments compared to transformer-CNN (80%) and UB-Net (79%). The improved performance can be attributed to the effective integration of local spatial features (CNN), global structural constraints (GAT), and robust non-linear regression (GBDT), which collectively enhance the method's ability to capture complex geological patterns while maintaining computational efficiency.

The proposed framework demonstrated competitive computational performance despite its architectural complexity. The end-to-end training time for the complete model was 52 min, compared to 45 min for transformer-CNN and 48 min for UB-Net. During inference, our method processed seismic data at 1.2 km^2/min, comparable to transformer-CNN (1.3 km^2/min) and significantly faster than UB-Net (0.9 km^2/min) due to the latter's Bayesian sampling requirements.

These results demonstrate that the proposed framework not only improves inversion accuracy but also provides practical

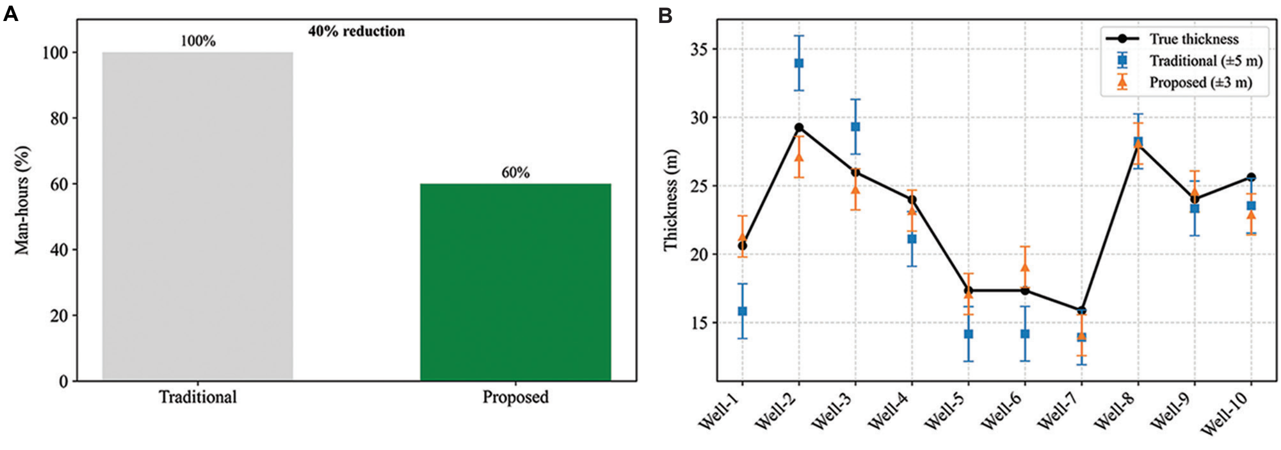


Figure 9. Industrial application effect verification. (A) Interpretation time reduction. (B) Reservoir thickness prediction.

solutions for industrial-scale seismic impedance inversion applications. The method's ability to generate high-quality 2D profiles and 3D cubes from conventional post-stack data makes it particularly valuable for reservoir characterization and geological interpretation in exploration projects.

Actual deployment tests showed that in a 200 km^2 3D work area, the combined approach reduced manual interpretation time by approximately 40%, while keeping the reservoir thickness prediction error within ± 3 m (Figure 9).

Figure 9A shows that the combined approach, by automatically generating high-precision impedance volumes, reduced geological interpretation time from an average of 1,200 man-hours to 720 man-hours (40% reduction). The efficiency improvement was particularly significant in fault interpretation (from 380 to 190 man-hours). As shown in Figure 9B, for thin reservoir predictions of 6–8 m thick, the combined approach achieved a relative error of only 8.3 ± 2.7%, compared to 22.5 ± 6.8% for traditional seismic inversion methods. Notably, this breakthrough was achieved while maintaining computational efficiency (processing time for a single work area was < 8 h), and its inference speed (1.2 km^2/min) fully meets the requirements of industrial production.

5. Discussion

The combined CNN–GAT–GBDT framework proposed in this study improved the accuracy of seismic impedance inversion, as evidenced by a 15.2% reduction in RMSE of seismic impedance inversion, particularly in complex geological environments. Traditional seismic impedance inversion methods are often affected by noise and geological complexity, resulting in significant deviations in inversion results. By combining the strengths of deep learning and GNNs, the proposed combined framework can adaptively extract multi-scale features from seismic data and effectively capture spatial dependencies among seismic traces through a GAT. This framework demonstrated advantages over traditional methods, with a 39.0% lower RMSE than SPI and a 22.7% lower RMSE than XGBoost in complex salt dome areas, particularly in areas with heterogeneity and complex structures.

However, despite the framework's impressive performance, several challenges remain, and there is room for improvement. First, while the combined framework effectively mitigated the impact of noise and outliers, its robustness to extreme outliers and noise still requires improvement. In practical geological exploration, some areas may experience significant geological changes or poor data quality. Improving the model's adaptability and reducing inversion errors under these extreme conditions remain urgent challenges. Second, although the framework utilized a joint loss function and end-to-end fine-tuning strategy during training, the computational and time costs of training also increase with larger geological dataset sizes. Therefore, further exploration is needed to optimize the model's computational efficiency, improve training speed, and enhance inference performance, particularly in practical applications of large-scale 3D seismic data.

In practice, the proposed CNN–GAT–GBDT framework provides an efficient and accurate solution for seismic impedance inversion, demonstrating significant advantages in industrial applications. By automatically generating high-precision impedance volumes, the framework significantly reduces geological interpretation time and achieves precise error control in reservoir thickness prediction. This performance improvement, particularly in predicting complex faults and thin reservoirs, demonstrates the potential of the framework in seismic exploration and is expected to be widely applied to various geophysical inversion tasks, such as lithology identification and reservoir characterization. Although the current framework has achieved some results, with continued technological advancement and application across a wider range of geological scenarios, further optimization and expansion of the framework will bring even greater breakthroughs in seismic data processing.

Overall, the combined application of CNNs, GATs, and GBDTs provides an innovative approach to seismic impedance inversion and advances the development of geophysical inversion technology. With the continued advancement of deep learning, GNNs, and ensemble learning techniques, this framework is expected to demonstrate its powerful application value in a wider range of seismic exploration tasks in the future.

6. Conclusion

In this study, a combined CNN–GAT–GBDT framework provided a novel approach to seismic impedance inversion, achieving a lower prediction error (RMSE of 285 m/s·g/cm^3) and higher structural similarity (SSIM of 0.93) compared to baseline models. By combining the strengths of deep learning and ensemble learning, this framework leverages the powerful spatial feature extraction capabilities of CNN, the unique strength of GAT in capturing non-Euclidean spatial relationships in seismic data, and the robust performance of GBDT in high-precision regression. This framework not only accurately captures detailed information in seismic data but also effectively suppresses noise interference, improving inversion accuracy. The reduction in error was more pronounced (22.7%) in complex geological conditions.

From a theoretical perspective, the innovation of the CNN–GAT–GBDT framework lies in its interdisciplinary integration of deep learning and GNNs. CNNs effectively extract local spatial features from seismic data through multi-scale convolutional operations, while GATs capture global dependencies of geological structures by modeling non-Euclidean relationships between seismic traces. The GBDT module integrates these deep features through ensemble learning and optimizes model predictions through non-linear regression, further improving prediction accuracy and robustness. This approach not only extracts rich spatial and spectral features from seismic data but also adapts to diverse geological scenarios in practical applications, demonstrating its powerful adaptive learning capabilities.

In terms of practical value, the CNN–GAT–GBDT framework provides an efficient and accurate solution for seismic impedance inversion, significantly improving inversion accuracy and computational efficiency, especially in complex geological environments. Experimental results demonstrate that the proposed framework achieved high-precision impedance inversion in a shorter time compared to traditional SPI methods and single deep learning models, and demonstrated high robustness and reliability in real-world applications. Furthermore, the forecast uncertainty estimates provided by the framework offer a quantitative basis for decision-making in geological exploration, significantly reducing risk and improving decision-making efficiency.

The innovative interdisciplinary approach makes the application of this framework in geophysical inversion profoundly significant. By combining deep learning techniques from computer science with GNNs from geology, it not only advances the development of seismic data processing technology but also provides new solutions to geophysical problems, such as seismic impedance inversion. In the future, this framework will not be limited to impedance inversion but can also be extended to other geophysical inversion tasks, such as lithology prediction and reservoir identification, showing broad application prospects.

Acknowledgments

None.

Funding

This research was financially supported by the Scientific Research Fund of Institute of Seismology, China Earthquake Administration and National Institute of Natural Hazards, MEM, (No. IS202226322); 2025 Doctoral Special Support Program Project of Chengdu Jincheng College (NO.2025JCKY(B)0018); the Key Research Base of Humanities and Social Sciences of the Education Department of Sichuan Province, Panzhihua University, Resource based City Development Research Center Project (NO.ZYZX-YB-2404); Mahasarakham University; and the Open Fund of Sichuan Oil and Gas Development Research Center (NO.2024SY017).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Tianwen Zhao, Guoqing Chen, Cong Pang, Palakorn Seenoi, Yiru Du

Formal analysis: Tianwen Zhao, Guoqing Chen, Cong Pang, Palakorn Seenoi, Nipada Papukdee

Investigation: Tianwen Zhao, Cong Pang, Piyapatr Busababodhin, Palakorn Seenoi, Nipada Papukdee, Yiru Du

Methodology: Tianwen Zhao, Guoqing Chen, Piyapatr Busababodhin, Palakorn Seenoi, Nipada Papukdee, Yiru Du

Validation: Tianwen Zhao, Guoqing Chen, Yiru Du

Writing–original draft: Tianwen Zhao, Guoqing Chen, Piyapatr Busababodhin, Nipada Papukdee, Yiru Du

Writing–review & editing: Tianwen Zhao, Guoqing Chen, Cong Pang, Piyapatr Busababodhin, Palakorn Seenoi, Yiru Du

Availability of data

Some data used in this study cannot be shared publicly due to collaborative agreement restrictions, but are available from the corresponding author on reasonable request.

References

1. Falade AO, Amigun JO, Abiola O. Hydrocarbon prospective study using seismic inversion and rock physics in an offshore field, Niger Delta. *Discov Geosci*. 2024;2(1):24.

   doi: 10.1007/s44288-024-00030-4

2. Zhang ZX, Gong F, Kozlovskaya E, Aladejare A. Characteristic impedance and its applications to rock and mining engineering. *Rock Mech Rock Eng*. 2023;56(4): 3139-3158.

   doi: 10.1007/s00603-023-03216-3

3. Su Y, Cao D, Liu S, Hou Z, Feng J. Seismic impedance inversion based on deep learning with geophysical constraints. *Geoenergy Sci Eng*. 2023;225:211671.

   doi: 10.1016/j.geoen.2023.211671

4. Lin Y. Multi-scale seismic impedance inversion based on Transformer model and deep learning. *Eng Res Express*.

2025;7(1):015209.

doi: 10.1088/2631-8695/ada48d

5. Wu X, Yan S, Bi Z, Zhang S, Si H. Deep learning for multidimensional seismic impedance inversion. *Geophysics*. 2021;86(5):R735-R745.

   doi: 10.1190/geo2020-0564.1

6. Li M, Yan XS, Zhang MZ. A comprehensive review of seismic inversion based on neural networks. *Earth Sci Inform*. 2023;16(4):2991-3021.

   doi: 10.1007/s12145-023-01079-4

7. Akingboye AS. Electrical and seismic refraction methods: Fundamental concepts, current trends, and emerging machine learning prospects. *Discov Geosci*. 2025;3(1):87.

   doi: 10.1007/s44288-025-00169-8

8. Leite EP, Vidal AC. 3D porosity prediction from seismic inversion and neural networks. *Comput Geosci*. 2011;37(8):1174-1180.

   doi: 10.1016/j.cageo.2010.08.001

9. Wallick BP, Giroldi L. Interpretation of full-azimuth broadband land data from Saudi Arabia and implications for improved inversion, reservoir characterization, and exploration. *Interpretation*. 2013;1(2):T167-T176.

   doi: 10.1190/INT-2013-0065.1

10. Chen H, Innanen KA, Chen T. Estimating P- and S-wave inverse quality factors from observed seismic data using an attenuative elastic impedance. *Geophysics*. 2018;83(2):R173-R187.

    doi: 10.1190/geo2017-0183.1

11. Okeugo CG, Onuoha KM, Ekwe CA, Anyiam OA, Dim CIP. Application of crossplot and prestack seismic-based impedance inversion for discrimination of lithofacies and fluid prediction in an old producing field, Eastern Niger Delta Basin. *J Pet Explor Prod Technol*. 2019;9(1):97-110.

    doi: 10.1007/s13202-018-0508-6

12. Azevedo L, Demyanov V. Multiscale uncertainty assessment in geostatistical seismic inversion. *Geophysics*. 2019;84(3):R355-R369.

    doi: 10.1190/geo2018-0329.1

13. Dai R, Yin C, Zaman N, Zhang F. Seismic inversion with adaptive edge-preserving smoothing preconditioning on impedance model. *Geophysics*. 2019;84(1):R11-R19.

    doi: 10.1190/geo2016-0672.1

14. Thibodeaux B, Ramsay T, Segovia F, Hernandez L, Ibrahim M. Closed-Loop Integrated Time-Lapse Seismic Feasibility in Amberjack Field–Deepwater Offshore Gulf of Mexico. In: *Paper Presented at: SPE Reservoir Characterization and Simulation Conference and Exhibition*. Dayeh University, Delta, Syria. SPE-196670-MS; 2019.

    doi: 10.2118/196670-MS

15. Zhang J, Li J, Chen X, Li Y. Geological structure-guided hybrid MCMC and Bayesian linearized inversion methodology. *J Pet Sci Eng*. 2021;199:108296.

    doi: 10.1016/j.petrol.2020.108296

16. Ma Q, Wang Y, Ao Y, Wang Q, Lu W. UB-Net: Improved seismic inversion based on uncertainty backpropagation. *IEEE Trans Geosci Remote Sens*. 2022;60:1-11.

    doi: 10.1109/TGRS.2022.3174911

17. Ning C, Wu B, Wu B. Transformer and convolutional hybrid neural network for seismic impedance inversion. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2024;17:4436-4449.

    doi: 10.1109/JSTARS.2024.3358610

18. Xiong W, Ji X, Ma Y, *et al.* Seismic fault detection with convolutional neural network. *Geophysics*. 2018;83(5): O97-O103.

    doi: 10.1190/geo2017-0666.1

19. An Y, Guo J, Ye Q, *et al*. Deep convolutional neural network for automatic fault recognition from 3D seismic datasets. *Comput Geosci*. 2021;153:104776.

    doi: 10.1016/j.cageo.2021.104776

20. Cao C, Wang X, Yang F, *et al*. Attention-driven graph convolutional neural networks for mineral prospectivity mapping. *Ore Geol Rev*. 2025;106554.

    doi: 10.1016/j.oregeorev.2025.106554

21. Yao G, Zhang Q, Zhang H, Li Y. Non-local self-similarity guided graph attention network for DAS-VSP noise and signal separation. *J Appl Geophys*. 2025;241:105835.

    doi: 10.1016/j.jappgeo.2025.105835

22. Zhou J, Gao Y, Lu J, Yin C, Han H. An ensemble learning algorithm for machinery fault diagnosis based on convolutional neural network and gradient boosting decision tree. *J Phys Conf Ser*. 2021;2025(1):012041.

    doi: 10.1088/1742-6596/2025/1/012041

23. Qian S, Peng T, Tao Z, *et al*. An evolutionary deep learning model based on XGBoost feature selection and Gaussian data augmentation for AQI prediction. *Process Saf Environ Prot*. 2024;191:836-851.

    doi: 10.1016/j.psep.2024.08.119

24. Li Q, Luo Y. High-resolution Bayesian sequential impedance inversion. In: *Paper Presented at: SEG International Exposition and Annual Meeting*; 2020.

    doi: 10.1093/jge/gxac035

25. Zhao T, Chen G, Suraphee S, Phoophiwfa T, Busababodhin P. A hybrid TCN-XGBoost model for agricultural product market price forecasting. *PLoS One*. 2025;20(5):e0322496.

    doi: 10.1371/journal.pone.0322496

26. Zhao T, Chen G, Pang C, Busababodhin P. Application and performance optimization of SLHS-TCN-XGBoost model in power demand forecasting. *Comput Model Eng Sci*. 2025;143(3):2883-2917.

    doi: 10.32604/cmes.2025.066442

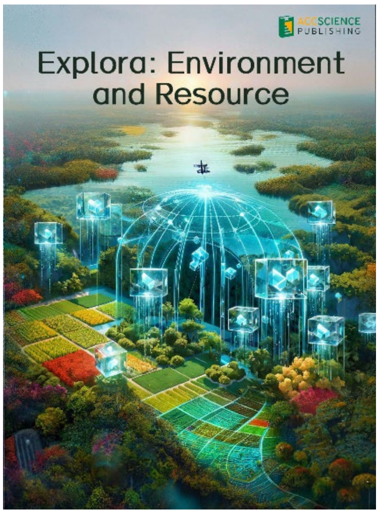
27. Bakurov I, Buzzelli M, Schettini R, Castelli M, Vanneschi L. Structural similarity index (SSIM) revisited: A data-driven approach. *Expert Syst Appl*. 2022;189:116087.

    doi: 10.1016/j.eswa.2021.116087

28. Brunet D, Vrscay ER, Wang Z. On the mathematical properties of the structural similarity index. *IEEE Trans Image Process*. 2011;21(4):1488-1499.

    doi: 10.1109/TIP.2011.2173206

OUR JOURNALS

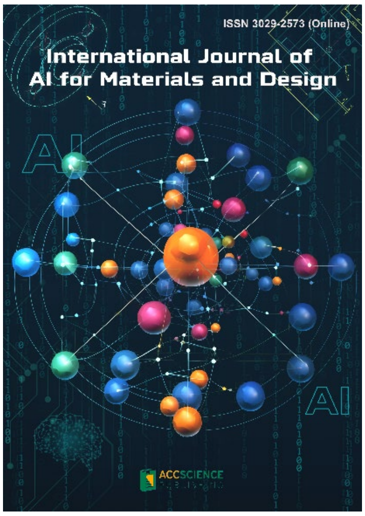


Explora: Environment and Resource (EER) is an international and multidisciplinary journal covering all aspects of the environmental impacts of socio-economic development. It is concerned with the complex interactions among society, development, and the environment, aiming to explore ways and means of achieving sustainability in all human activities related to development.

EER covers subject areas, including but not limited to the following:

- Water reclamation, wastewater treatment and waste management for eco-environment sustainability
- Innovative nanotechnology, catalysis, photocatalysis and nano materials for decontamination practices and renewable energy resources.
- Climate change-related sustainable resource and environment quality analysis
- Microorganisms and plants-oriented low-carbon processes and strategies for clean energy, water and resource management

International Journal of AI for Materials and Design is an international, peer-reviewed open-access journal that aims to bridge the cutting-edge research between AI and materials, AI and design. In recent years, the tremendous progress in AI is leading a radical shift of AI research from a mainly academic endeavor to a much broader field with increasing industrial and governmental investments. The maturation of AI technology brings about a step change in the scientific research of various domains, especially in the world of materials and design. Machine learning (ML) algorithms enable researchers to analyze extensive datasets on material properties and accurately predict their behavior in different conditions. This subsequently impact the industry to leverage on big data and advanced analytics to build scientific strategies, scale operational performance of processes and drive innovation. In addition, AI and ML are uniquely positioned to enable advanced manufacturing technologies across the value chain of different industries. Integration of multiple and complementary AI techniques, such as ML, search, reasoning, planning, and knowledge representation, will further accelerate advances in scientific discoveries, engineering excellence and the future of cyber-physical systems manufacturing.



International Journal of AI for Materials and Design covers the following topics: AI or machine learning for material discovery, AI for process optimization, AI and data-driven approaches for product or systems design, application of AI in advanced manufacturing processes such as additive manufacturing, IoT, sensors, robotics, cloud-based manufacturing, intelligent manufacturing for various applications, autonomous experiments, material intelligence, energy intelligence, and AI-linked decarbonization technologies.

Start a new journal

Write to us via email if you are interested to start a new journal with AccScience Publishing. Please attach your CV, professional profile page and a brief pitch proposal in your email. We shall inform you of our decision whether we are interested to collaborate in starting a new journal.

Contact: info@accscience.com

https://accscience.com/journal/JSE

ACCSCIENCE
PUBLISHING

Access Science Without Barriers

Contact

www.accscience.com
9 Raffles Place, Republic Plaza 1 #06-00 Singapore 048619
E-mail: editorial@accscience.com
Phone: +65 8182 1586