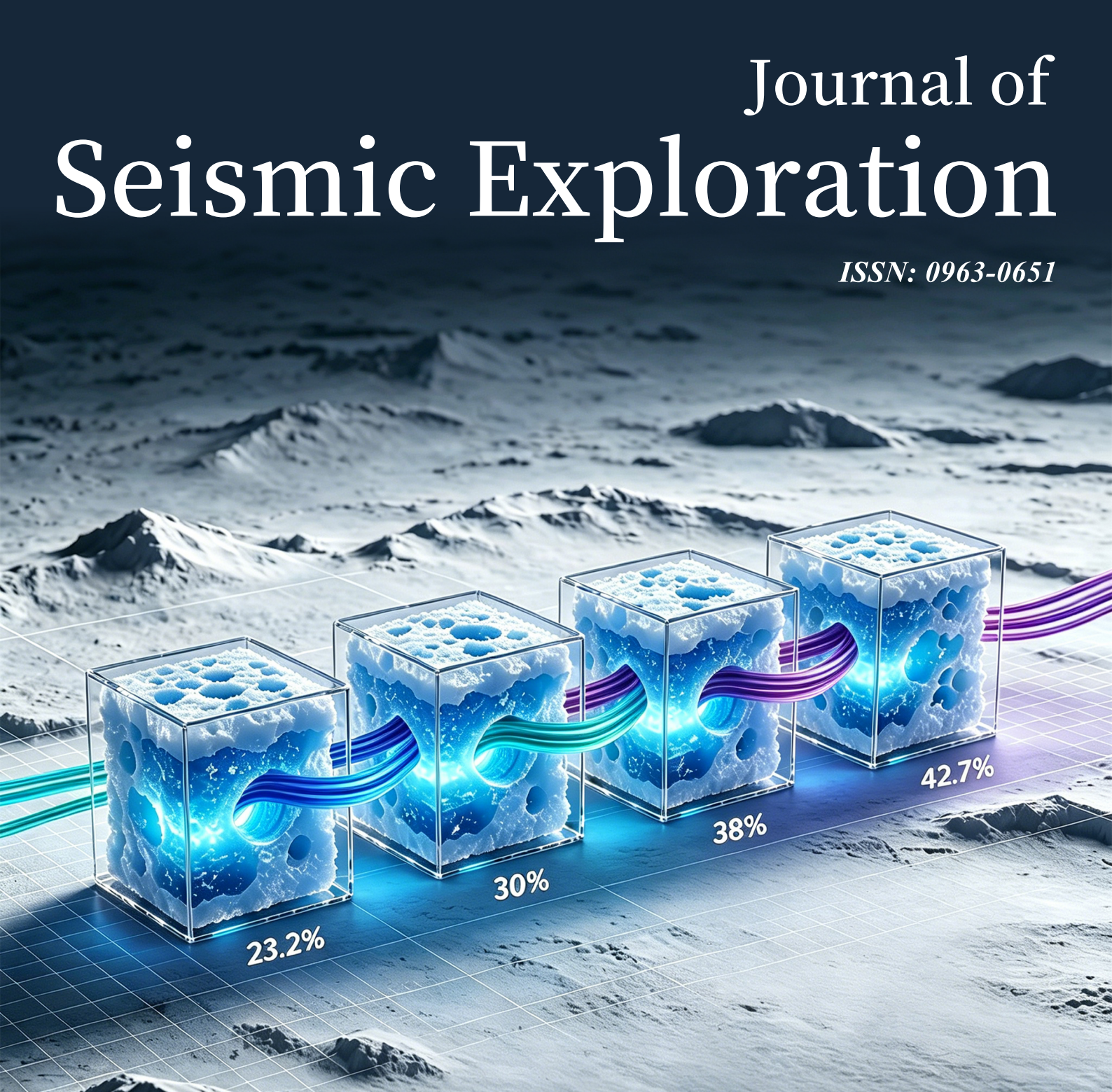
Journal of
# Seismic Exploration

ISSN: 0963-0651

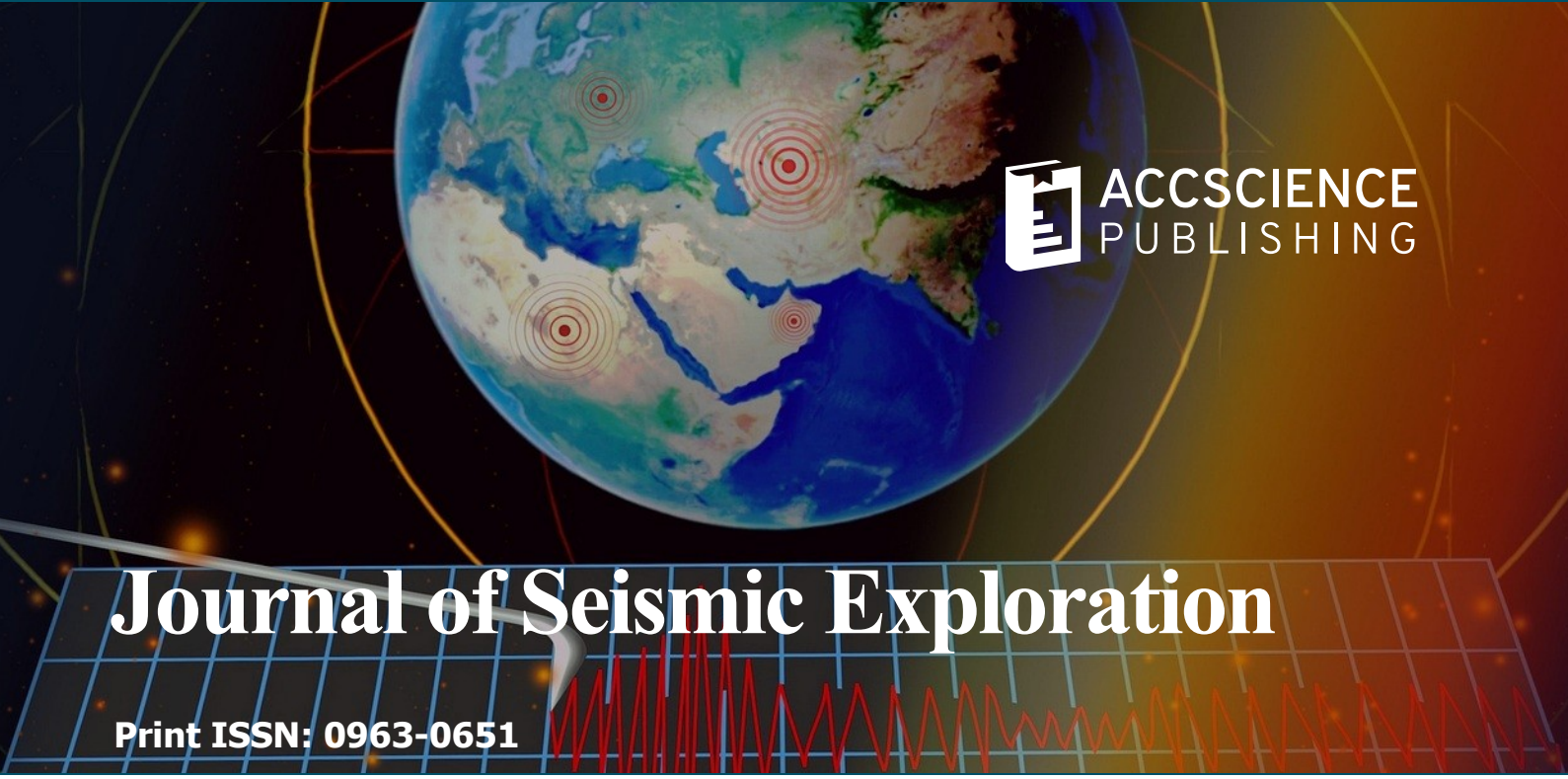



## Elastic modulus analysis of a 3D-printed permafrost simulation model with varying porosity in a seismic physical modeling

Volume 35 · Issue 2
April 2026

ACCSCIENCE PUBLISHING

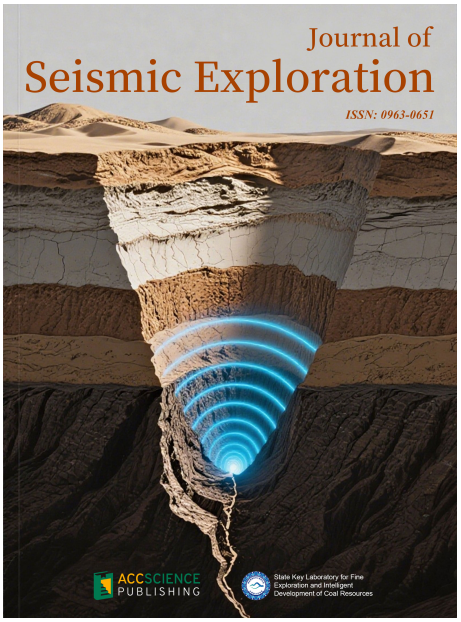
State Key Laboratory for Fine Exploration and Intelligent Development of Coal Resources



# 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.




## 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
2 Venture Drive, #07-06 Vision Exchange, Singapore 608526

Volume 35 • Issue 2 • 2026
ISSN 0963-0651 (print)

# Journal of Seismic Exploration

**Editors-in-Chief**

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

**Arthur B. Weglein**
*University of Houston*

ACCSCIENCE PUBLISHING

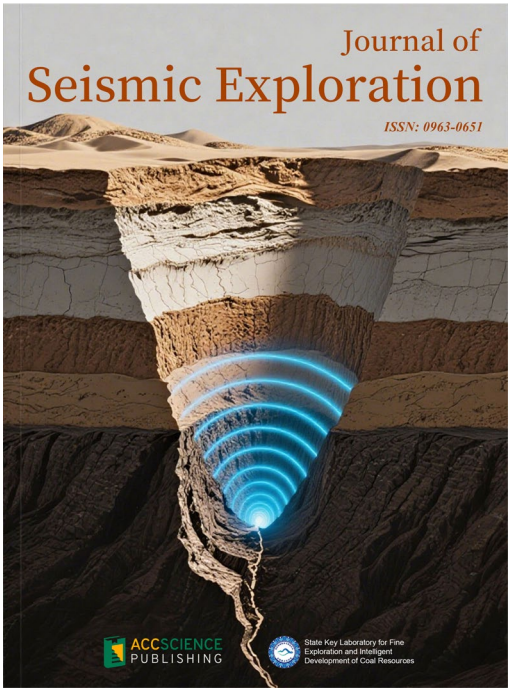
Access Science Without Barriers

**Full issue copyright © 2026 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/35/2.

**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

Suping Peng
Arthur B. Weglein

## Executive Editor-in-Chief

Jing Ba

## Editorial Board Members

Mauricio Sacchi
Ilya Tsvankin
Li-Yun Fu
Boris Gurevich
José M. Carcione
Haijiang Zhang
Jun Matsushima
Fabio Trippetta
Qizhen Du
Lasse Amundsen
Jingyi Chen
Yangkang Chen
Lin Chen
Vladimir Cheverda
Eduardo Filpo
Doug Foster
Bill N. Goodway
Songfeng Guo
Zhiqi Guo
Cory Hoelting
Morten Jakobsen
Gilles Lambaré
Evgeny Landa
Jing Li
Huailiang Li
Chih-Ping Lin
José Eduardo M. Lira
Yang Liu
Jianwei Ma
George McMechan
Mark A. Meier
Roman Pevzner
Flavio Poletto
J. Germán Rubino
Mrinal K. Sen
Serge A. Shapiro
Changsoo Shin
Juan I. Soto
Paul L. Stoffa
Walter Söllner
Dinghui Yang
Umberta Tinivella
Peter Traynin
Arie Verdel
Dirk Jacob Eric Verschuur
Xiujuan Wang
Yibo Wang
Jincai Zhang
Lele Zhang
Gulan Zhang
Wei Zhang
Huai Zhang
Fons ten Kroode

## Early Career Editorial Board

Cong Luo
Xintong Dong
Fei Gong
Qiang Guo
Guangtan Huang
Zhongxiao Li
Xingye Liu
Shaoyong Liu
Xinpeng Pan
Yingming Qu
Xiaokai Wang
Haibo Wu
Pan Zhang
Sheng Zhang
Jidong Yang
Peng Lin
Wei Chen
Yong Hu
Yaju Hao
Jiachun You

ACCSCIENCE PUBLISHING

April 2026 | Volume 35 | Issue 2 |

Journal of Seismic Exploration

https://accscience.com/journal/JSE

# CONTENTS

**ARTICLES**

**1** **Random noise suppression in node seismometer signals using improved complete ensemble empirical mode decomposition with adaptive noise and grey relational analysis**
*Cong Pang, Tianwen Zhao, Guoqing Chen, Yuxuan Liang, Xingxing Li, Ya Xiang, Sirui Liu, Piyapatr Busababodhin*

**31** **Seismic azimuthal amplitude variation with offset response analysis in coalbed methane reservoirs with aligned fractures using Biot theory**
*ZhaoJi Zhang, Fei Gong, GuanGui Zou, Qiang Guo, GuoWei Zhu, Hao Li*

**50** **Mitigating numerical dispersion in full-waveform inversion imaging: A preconditioned optimization approach for finite-difference weights**
*Ganglin Lei, Jianping Huang, Wensheng Duan, Chao Chen, Zhenwen Liu, Chang Zhou, Weiting Peng*

**65** **Elastic modulus analysis of a 3D-printed permafrost simulation model with varying porosity in a seismic physical modeling**
*Daechul Kim, Yeonjin Choi, Wookeen Chung, Sungryul Shin*

**78** **Seismic characterization and spatial distribution analysis of coal gangue in the deep coalbed of the Northeastern Ordos Basin**
*Zelei Jiang, Xuri Huang, Dong Zhang, Yucong Huang, Yong Wu*

**93** **TFDenoiser-Edge: A hybrid convolutional neural network–Transformer framework for real-time seismic denoising on edge devices in extreme environments**
*Zesheng Yang, Qingfeng Xue, Xiaoning Wang, Tao Wang*

**108** **Critical metal enrichment in coal-bearing strata of the eastern Ordos Basin: A geophysical and geochemical study**
*Yi Yang, Jingtao Zhao, Suzhen Shi, Wanli Gao*

**138** **Automatic differentiation-based weighted instantaneous phase inversion**
*Mengzi Wang, Yong Hu, Zhihan Zhang, Chunyuan Shi, Xiaoyuan Liao*

**151** **Direct envelope-based reflection full waveform inversion for shallow marine seismic data**
*Kyoungmin Lim, Wookeen Chung, Jungkyun Shin, Jiho Ha*

**168** **Enhanced amplitude variation with incident angle inversion with Cauchy regularization using ensemble smoother-based data assimilation**
*Rui Liao, Ronghuo Dai*

**181** **Advances in intelligent seismic interpretation for onshore unconventional reservoirs in CNOOC: The Linxing–Shenfu gas field case**
*Wenlan Li, Qixin Li, Xiaowen Zheng, Bo Wang, Di Wang, Jicai Ding*

**201** **XGBoost-Deep Residual (FedXGB-ResNet) collaborative porosity prediction in a federated learning framework**
*Tianwen Zhao, Guoqing Chen, Junyan Li, Cong Pang, Palakorn Seenoi, Nipada Papukdee, Piyapatr Busababodhin*

225 **Computational efficiency of deep learning-based seismic fault interpretation considering context window and input resolution**
*JBowen Deng, Guangui Zou, Suping Peng, Chengyang Han, Jingwen Xue*

241 **GeoSSL: A geology-aware self-supervised framework for fault detection in 3D seismic data**
*Mengge Wang, Xinrong Hu, Junyu Zhang, Haofa Lin*

271 **A two-branch convolutional neural network for pre-stack elastic parameter inversion of coal-bearing gas reservoirs**
*Fei Li, Mengbo Zhang, Qiang Liang, Xiaojie Cui, Na Ni, Qingzhou Zhang, Yongheng Zhang*

290 **Fracture prediction based on W-transform inversion spectral decomposition**
*Xiuwei Wang, Ying Jia, Xilin Qin, Jun Wang, Yinglong Li*

310 **Interpretable intelligent gas-bearing reservoir prediction using time–frequency analysis and manifold-regularized semi-supervised GAN**
*Shuying Ma, Junxing Cao, Rong Wang, Xudong Jiang, Jun Wang, Lingsen Zhao, Hong Li, Xin Tang*

331 **Two-way acoustic wave-equation-based wavefield depth extrapolation for imaging in vertically transversely isotropic media**
*Shanzheng Hu, Pengyuan Sun, Shuqin Li, Feng Hu, Min Guan, Jiachun You*

354 **Suppression of surface waves in seismic data using an adaptive time–frequency–wavenumber filter**
*Xuejie Gao, Yong Wang, Youjuan He, Zhili Chen, Zerun Nian, Lang Yang*

371 **Reducing interpretation risks in organic-rich low-velocity mudstone using rock physics modeling in Yinggehai Basin, South China Sea**
*Yanhua Xie, Jinpeng Li, Fang Li, Wanyuan Sun, Yayun Dang*

391 **Seafloor reflector imaging in 2D seismic data through muting of out-of-plane signals in the Ulleung Basin, East Sea**
*Ganghoon Lee, Changyoon Lee, Junseok Kwon, Snons Cheong*

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Random noise suppression in node seismometer signals using improved complete ensemble empirical mode decomposition with adaptive noise and grey relational analysis

**Cong Pang[1,2], Tianwen Zhao[3], Guoqing Chen[4,5], Yuxuan Liang[4], Xingxing Li[1], Ya Xiang[1], Sirui Liu[1], and Piyapatr Busababodhin[5]***

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

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

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

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

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

## Abstract

Node seismometer signals are often contaminated with substantial environmental noise due to the complex conditions encountered in geophysical exploration and seismic monitoring. This necessitates high precision in the pre-processing of ground motion signals, as inadequate processing may compromise subsequent operations, such as P-wave first-arrival time extraction, peak energy calculation, ground motion period determination, and magnitude estimation. To obtain more authentic seismic waveforms, a node seismometer signal denoising model is proposed. This model integrates grey relational analysis (GRA) with improved complete ensemble empirical modal decomposition adaptive noise (ICEEMDAN). This method first decomposes the noisy signal using ICEEMDAN to obtain multiple intrinsic mode functions (IMFs), which are then sequentially arranged and labeled. Subsequently, for each IMF, the correlation coefficient, mutual information, $R^2$, adjusted $R^2$, Jensen–Shannon divergence, cosine similarity, root mean squared error, mean absolute error, mean absolute percentage error, and sample entropy were calculated, forming an evaluation matrix for assessing the reliability of all IMFs. Finally, using GRA, the correlation coefficients and degrees of association between each evaluation metric and different IMF components were calculated. The IMF components were ranked based on their association degrees to determine the relative effectiveness of their signal components. Linear reconstruction on the top-ranked IMF components was performed to complete the signal denoising process proposed. Experiments on denoising simulated seismic signals, recorded seismic event signals, and recorded ground motion signals all demonstrate that the GRA–ICEEMDAN model outperforms classical denoising methods. The comprehensive denoising scores for the three experiments were 100, 98.0180, and 93.9056, respectively, with signal-to-noise ratio improvements reaching 24.0049 dB, 20.8926 dB, and 16.3523 dB, respectively. The model effectively distinguishes noise components from effective components, with minimal reconstruction errors

***Corresponding author:**
Piyapatr Busababodhin
(65010263001@msu.ac.th)

**Citation:** Pang C, Zhao T, Chen G, *et al.* Random noise suppression in node seismometer signals using improved complete ensemble empirical mode decomposition with adaptive noise and grey relational analysis. *J Seismic Explor.* 2026;35(2):025440098.
doi: 10.36922/JSE025440098

**Received:** October 30, 2025

**Revised:** December 22, 2025

**Accepted:** December 29, 2025

**Published online:** March 4, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

and signal loss after original signal decomposition, making it suitable for seismic monitoring and geophysical exploration involving small-to-medium sample sizes.

***Keywords:*** Node seismometer; Random noise suppression; Improved complete ensemble empirical modal decomposition adaptive noise; Grey relational analysis; Seismic signal denoising; Ground pulsation signal denoising

## 1. Introduction

As exploration tasks become increasingly complex and exploration accuracy requirements continue to increase, node seismometers, as a new type of seismic data acquisition equipment, have gradually become an important technical means for high-density seismic acquisition due to their cable-free design, flexible deployment, and adaptability to complex environments.[1-3] Compared with traditional wired seismometers, node seismometers have demonstrated significant advantages in complex surface areas (such as mountain canyons, swamps and lakes, and urban underground spaces), effectively overcoming the deployment limitations of wired systems and improving acquisition efficiency. However, node seismometers still face severe challenges in practical applications. Due to the use of wireless data transmission and the complex working environment, the acquired seismic signals are more susceptible to environmental noise, instrument electronic noise, and random interference, resulting in signal quality degradation.[4-6]

Given the increasing use of velocity-type node seismic instruments in mobile seismic activity monitoring and underground active fault detection, there is a strong practical need to study data pre-processing methods, such as denoising. By reducing the random environmental noise content in the original signal, the quality of seismic data can be further improved, thereby enhancing the accuracy of extracting parameters, such as the first arrival time of the first wave.

The core of seismic signal denoising algorithms lies in completely suppressing noise while maintaining the waveform characteristics and dynamic characteristics of the effective signal to avoid distortion. Seismic signal denoising methods have evolved from traditional Fourier analysis to modern time–frequency analysis. Early methods were mainly based on the frequency domain characteristics of the signal, such as the Fourier transform and bandpass filtering. These methods assumed that the signal and noise were separable in the frequency domain; exhibiting good results for stationary signal processing.[7-11] However, actual seismic signals have the characteristics of non-stationary and non-minimum phase. Traditional frequency domain methods often damage the effective signal while denoising, especially the phase information of waveform mutations.

Wavelet transform, as an important tool for time–frequency analysis, realizes multi-scale analysis of non-stationary signals through adjustable time–frequency windows and has been widely used in the field of seismic denoising.[12] The threshold denoising method based on wavelet transform achieves signal–noise separation by setting an appropriate threshold according to the different performances of signal and noise in the wavelet domain, suppressing high-frequency noise while enhancing the ability to retain low-frequency effective signals. However, the effect of the wavelet transform depends largely on the selection of the wavelet basis function and decomposition scale, and its adaptability to transient components in non-stationary signals is limited.

To overcome the shortcomings of the wavelet transform, the empirical mode decomposition (EMD) method was introduced into the field of signal processing.[13] EMD performs adaptive decomposition based on the local characteristic time scale of the signal, decomposing complex signals into a series of intrinsic mode functions (IMFs), which are very suitable for processing non-stationary and non-linear signals. However, traditional EMD faces challenges, such as modal aliasing and endpoint effects, which affect the stability and reliability of the decomposition. To address the shortcomings of EMD, researchers have proposed a series of improved algorithms, such as ensemble EMD (EEMD), complete EEMD (CEEMD), and CEEMD with adaptive noise (CEEMDAN). These algorithms reduce modal mixing and reconstruction errors by introducing noise-assisted analysis.[14-16] Furthermore, improved CEEMDAN (ICEEMDAN) optimizes the noise addition method based on CEEMD. Instead of adding random white noise to the original signal, it performs EMD by introducing a pair of white noises with opposite properties. This can suppress the reconstruction error caused by the introduction of white noise to a certain extent and control the amount of residual noise, resulting in a final decomposition that is more accurate.[17-20]

Denoising methods based on signal decomposition can subdivide the original noisy signal into IMFs with high

noise content, low noise content, and near-noise content. However, after decomposition, common problems exist, such as unclear discrimination of the noise content of IMFs, monotonous and easily discarded IMF waveforms containing the main amplitude energy, and difficulty in selecting the reconstruction targets. An appropriate evaluation index system has not been established to measure the quality or noise content of these IMFs. Hence, an effective evaluation method is urgently needed to accurately identify the effective IMF components that can participate in the final reconstruction and provide a reasonable basis for IMF screening. On the other hand, grey relational analysis (GRA), a mathematical method for measuring the degree of correlation between factors, has been introduced into the field of signal processing in recent years, providing a new criterion for distinguishing signal from noise.[21-23] GRA can adaptively identify the dominant noise component by calculating the geometric similarity between sequences and has demonstrated good performance in image denoising. However, research combining GRA with ICEEMDAN for node seismometer signal processing is still scarce. This technical integration is expected to achieve new breakthroughs in noise suppression and signal fidelity.

To fully utilize the advantages of ICEEMDAN in adaptive decomposition of non-stationary signals and combine the sensitivity of GRA in distinguishing between signals and noise, this study proposes a nodal seismograph signal denoising model based on GRA optimizing CEEMD to solve the problem of random noise interference in nodal seismograph-acquired signals. The model comprehensively considers the correlation coefficient (COEF),[24] mutual information (MI),[25] sample entropy (SE),[26] mean absolute error (MAE),[27] mean absolute percentage error (MAPE),[28] coefficient of determination (also known as $R^2$),[29] cosine similarity (CS),[30] root mean square error (RMSE),[31] Jensen–Shannon divergence JSD,[32] and adjusted $R^2$ (Adj-$R^2$).[33] GRA is then used to pe=rform multi-dimensional compression of the IMF characteristic indicator matrix, thereby constructing a comprehensive evaluation function of the quality of IMF components. Based on the ranking of this function, the original noisy signal is linearly reconstructed. Finally, simulation signal experiments and node seismometer measured signal denoising experiments are designed, and the proposed GRA–ICEEMDAN model is compared with classical denoising models, such as Kalman filtering, low-pass finite impulse response (FIR) filtering, and Savitzky–Golay, to evaluate overall denoising performance.

The main research contents and innovations of this study include:

(i) Application and optimization of the ICEEMDAN signal decomposition method: Using ICEEMDAN to decompose the node seismometer signal into a series of IMF components, optimize the decomposition parameters according to the characteristics of the seismic signal, improve the modal aliasing problem of the traditional EMD algorithm, and lay the foundation for subsequent noise separation.

(ii) Introduction of GRA criteria: GRA is innovatively applied to the classification and discrimination of IMF components. By calculating the grey correlation between each IMF and the noisy original signal, the noise-dominated IMF and the signal-dominated IMF are adaptively distinguished; overcoming the limitations of traditional methods based on COEFs or fixed thresholds.

(iii) Construction of hybrid denoising model: Combining ICEEMDAN with GRA establishes a complete denoising process, including signal adaptive decomposition, IMF component grey correlation calculation, noise component identification and screening, signal reconstruction, and other steps, that effectively suppresses random noise and preserves effective signals.

(iv) Algorithm verification and performance evaluation: Through simulation experiments, ground pulsation signal denoising experiments and earthquake event signal denoising experiments, the effectiveness of the proposed method is systematically verified, and compared with traditional and modern methods, such as Kalman filtering, low-pass FIR filtering, Savitzky–Golay filtering, and others, to evaluate its performance in terms of signal-to-noise ratio (SNR) improvement, RMSE reduction, and waveform fidelity.

(v) Design of comprehensive noise reduction evaluation metrics: To accurately assess the noise reduction capability and fidelity of various noise reduction models, five widely recognized and complementary quantitative evaluation metrics were selected from multiple perspectives, including signal fidelity, noise suppression intensity, and frequency domain feature preservation. These metrics were integrated to design a single comprehensive evaluation metric. The sub-metrics include: SNR improvement, noise suppression rate, correlation, frequency similarity, and energy retention rate. The overall score metric is defined as the sum of the normalized weighted values of the five indicators (with the first indicator in simulation experiments replaced by traditional SNR), where each indicator carries equal weight.

The structure of this study is arranged as follows: Section 1 is the introduction, which introduces the research

background and significance; Section 2 explores the GRA–ICEEMDAN signal denoising model, which elaborates on the denoising model principle and implementation steps of GRA combined with ICEEMDAN; Section 3 systematically analyzes the random noise suppression experiment and analysis of simulated noisy signals; Section 4 provides a report on denoising experiments and analysis of node seismometer recorded signals through two sets of event data with different environmental stimuli, and a comparative analysis of similar filtering models; and lastly, Section 5 summarizes the research results and proposes future research directions.

## 2. The GRA–improved complete ensemble EMD with adaptive noise signal denoising model

### 2.1. Analysis of the principles of GRA

GRA is a comprehensive evaluation method for small-sample problems based on the grey system theory and the similarity and difference of indicators to obtain a degree of correlation. It generally goes through several main steps, including data standardization, COEF calculation, correlation degree calculation, and ranking correlation degree. It has many advantages, such as small sample adaptability, no need for strict distribution assumptions, strong noise resistance, and low computational complexity. The general steps for performing correlation analysis between evaluation objects and indicators using GRA are as follows:

(i) Set the number of evaluation objects to $m$ and the number of evaluation indicators to $n$. Standardize the original data and set the reference sequence and comparison sequence as $x_0 = \{x_0(k)\ k = 1,2,\ldots,n\}, x_i = \{x_i(k) \mid k = 1,2,\ldots,n; i = 1,2,\ldots,m\}$, respectively.

(ii) Determine the corresponding weights of each evaluation indicator and construct a one-dimensional weight vector $w = [w_1, w_2, \ldots, w_n]$, where $w_k (k = 1,2,\ldots,n)$ represents the weight value of the $k$-th evaluation indicator.

(iii) Set the resolution coefficient $\rho \in [0,1]$. The grey COEF $\xi_i(k)$ between the comparison sequence $x_0$ and the reference sequence $x_i$ on the $k$-th evaluation indicator is defined as **Equation 1:**

$$\xi_i(k) = \frac{\min\limits_s \min\limits_t \left|x_0(t) - x_i(t)\right| + \rho \max\limits_s \max\limits_t \left|x_0(t) - x_i(t)\right|}{\left|x_0(k) - x_i(k)\right| + \rho \max\limits_s \max\limits_t \left|x_0(t) - x_i(t)\right|} \quad (1)$$

where, $\min\limits_s \min\limits_t \left|x_0(t) - x_i(t)\right|$ is the minimum difference between the two levels; $\max\limits_s \max\limits_t \left|x_0(t) - x_s(t)\right|$ is the maximum difference between the two levels.

(iv) According to the index weight vector $w$ and $m \times n$ grey COEFs $\xi_i(k)$, the grey weighted correlation degree of the $i$-th evaluation object is calculated as in **Equation 2:**

$$r_i = \sum_{k=1}^{n} w_i \xi_i(k) \quad (2)$$

(v) Finally, the gray weighted correlation degree obtained in step (iv) is used to score and rank the $m$ evaluation objects. The higher the correlation degree, the higher the credibility of the evaluation object.

### 2.2. The improved complete ensemble EMD with an adaptive noise signal denoising method

The ICEEMDAN method is a multi-step improvement on the standard EMD technique. Having undergone various stages of development and improvement, including EMD, EEMD, and CEEMDAN, ICEEMDAN introduces a local mean operator to replace direct modal decomposition and optimizes noise injection strategies based on the decomposition of non-stationary and non-linear signals. This overcomes common signal decomposition issues, such as modal aliasing, residual noise, and pseudo-modal components. EMD suffers from modal aliasing. EEMD introduces white noise to balance the original noise, but still leaves residual reconstruction errors and pseudo-components. CEEMDAN reduces reconstruction errors, but the early IMFs still contain residual noise, and signal information is delayed. The main steps for signal decomposition are:

(i) The original signal $x$ is decomposed to construct a noisy synthetic signal (**Equation 3**):

$$x^{(i)} = x + \beta_0 E_1(w^{(i)}) \quad (3)$$

where, $\beta_0$ is the signal-to-noise ratio at the first decomposition; $w^{(i)}$ is the $i$-th Gaussian white noise added; and $E_1(\cdot)$ is the first IMF component obtained by EMD decomposition.

(ii) Repeat the previous step $N$ times to obtain $N$ $x^{(i)}$ and calculate the average of all local mean function values, which is the first residual component, as shown in **Equation 4:**

$$r_1 = \frac{1}{N} \sum_{i=1}^{N} M(x^{(i)}) \quad (4)$$

In the above formula, $M(\cdot)$ is the local mean function.

(iii) The first mode is obtained by subtracting the first residual component from the original signal, as shown in **Equation 5:**

$$d_1 = x - r_1 \quad (5)$$

(iv) Based on the local mean function value calculated after the $k$-th decomposition, the $k$-th residual component is obtained, and the residual result of this decomposition is subtracted from the previous residual component to obtain the $k$-th mode $d_m$, as shown in **Equations 6** and **7**

$$d_m = r_{k-1} - r_k \tag{6}$$

$$r_k = \frac{1}{N}\sum_{i=1}^{N} M(r_{k-1} + \beta_{k-1} E_k(\omega^{(i)}) \tag{7}$$

(v) Continue executing step (iv) until the decomposition stopping condition is met.

As a data-driven adaptive decomposition method, the number of IMFs generated by ICEEMDAN is not fixed but depends on the time–frequency complexity of the input signal itself and the algorithm's stopping criteria (such as whether the algorithmic residual meets the condition of a monotonic function). Specifically, the richer the oscillatory modes within a signal, the greater the number of IMFs typically decomposed; conversely, fewer IMFs are generated. This adaptability is precisely where it surpasses fixed-basis transforms, such as the wavelet transform, enabling it to more flexibly match the non-stationary characteristics of seismic signals.

### 2.3. Signal decomposition evaluation metrics and noise reduction effect evaluation metrics

To measure the differences or similarities between each IMF obtained from the ICEEMDAN decomposition and the original signal, and identify the effective and noise components in the IMF component set, thereby achieving high-quality denoising of the original noisy signal, we selected 10 parameters: COEF, MI, SE, MAE, MAPE, $R^2$, CS, RMSE, JSD, and Adj-$R^2$. These parameters were used to evaluate the overall performance of all IMFs from multiple perspectives, including waveform similarity, amplitude error, and signal disorder.

The value of JS is between 0 and 1. The value 0 indicates identical distributions, and 1 indicates opposite distributions. JS divergence measures the similarity between two probability distributions. CS is a value between −1 and 1, where 1 indicates identical directions (two vectors are completely similar), and 0 indicates no similarity (two vectors are perpendicular and have completely different directions).

Sample entropy is a characteristic parameter describing the degree of non-stationarity of a signal. If the signal has no obvious distribution pattern, experiences severe local oscillations, or has a lot of noise, the sample entropy value is large, and vice versa. The calculation of $N$-dimensional signal sample entropy generally involves phase space reconstruction, calculation of reconstructed vector distance, calculation of matching degree, and its mean $B^m(r)$ under the similarity tolerance threshold $r$, increasing the dimension $m$ to $m$+1, and repeating the above process to obtain the sample entropy value.

In linear regression, Adj-$R^2$ is used to correct the spurious improvement in model fit caused by an increase in the number of independent variables. The calculation formulas for all evaluation indicators are as follows in **Equations 8-17**:

$$COEF = \frac{\sum_{i=1}^{n} s(i)y(i)}{\sqrt{\sum_{i=1}^{n} y^2(i)\sum_{i=1}^{n} s^2(i)}} \tag{8}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|s(i) - y(i)\right| \tag{9}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{s(i) - y(i)}{s(i)}\right| \tag{10}$$

$$CS = \frac{\sum_{n=1}^{N}\left(y(i)\cdot s(i)\right)}{\sqrt{\sum_{n=1}^{N} y^2(i)}\cdot\sqrt{\sum_{n=1}^{N} s^2(i)}} \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(s(i) - y(i)\right)^2} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(s(i) - y(i)\right)^2}{\sum_{i=1}^{n}\left(s(i) - \bar{s}\right)^2} \tag{13}$$

$$JSD = \frac{1}{2}\sum_{i}\left(s(i)\times log\left(2\cdot s(i)/(s(i)+y(i))\right)\right) + \frac{1}{2}\sum_{i}\left(y(i)\times log\left(2\cdot y(i)/(s(i)+y(i))\right)\right) \tag{14}$$

$$MI(Y;S) = \sum_{y\in Y}\sum_{s\in S} p(y,s)log\left(\frac{p(y,s)}{p(y)\cdot(s)}\right) \tag{15}$$

$$SampEn(m,r,N) = -ln\frac{B^{m+1}(r)}{B^m(r)} \tag{16}$$

$$Adj - R^2 = 1 - \frac{\left(1-R^2\right)\left(n-1\right)}{n-p-1} \tag{17}$$

In the above equations: $y(n)$ is the IMF component obtained by ICEEMDAN; $s(n)$ is the original signal; $\bar{s}(n)$ is the mean of the original signal, $n$ is the signal length, $p$ is

the number of independent variables; $p(s,y)$ is the joint probability distribution of discrete variables $S$ and $Y$; and $p(y)$ and $p(s)$ are the marginal probability distributions of discrete variables $S$ and $Y$, respectively.

To accurately evaluate the noise reduction capability and fidelity of various denoising models, five widely recognized and complementary quantitative evaluation metrics were selected from multiple perspectives, including signal fidelity, noise suppression intensity, and frequency domain feature preservation. An overall score was designed based on these metrics:

In practical applications lacking pure reference signals, SNR improvement serves as the most direct measure of noise suppression effectiveness. Assuming the signal's useful components are primarily distributed in the low-frequency range, while the high-frequency portion consists mainly of noise, the noise component is estimated separately in both the noisy signal and the denoised signal. The SNR is then calculated for each and compared. The difference between these two SNR values represents the final SNR improvement value, measured in dB. The noise estimation algorithm employs a high-pass filter to isolate high-frequency noise while utilizing a low-pass filter to estimate the signal components.

Assuming the useful signal components are concentrated in the mid-to-low frequency range, the noise suppression rate metric quantifies the degree of noise suppression by comparing the energy changes in the high-frequency portion of the signal before and after denoising. It is defined as the difference between the high-frequency energy ratio of the denoised signal and that of the original signal. This indicator directly quantifies the proportion of noise components removed.

Correlation calculates the Pearson COEF between the denoised signal and either a true clean signal (in simulations) or a noise reference signal (in real data). A high COEF (close to 1) indicates the denoising process effectively preserves the original signal's structure and waveform characteristics, avoiding excessive distortion.

Frequency similarity is defined as the COEF between the spectral value sequences of the pre- and post-denoising signals after the Fourier transform. This objectively evaluates the method's ability to preserve useful signal components in the frequency domain and suppress noise bands.

The denoising process should ensure useful signal energy is not excessively attenuated. Energy retention is defined as the ratio of the energy of the denoised signal (sum of squares at each point) to the energy of the original noisy signal, balancing noise suppression and signal preservation.

Overall score serves as a comprehensive evaluation metric for denoising effectiveness; it is defined as the sum of the normalized weighted values of the above five metrics (with the first metric in simulation experiments replaced by traditional SNR), where each metric carries equal weight.

### 2.4. Signal denoising process based on the GRA–improved complete ensemble EMD with adaptive noise signal denoising model

The signal processing flow of the GRA–ICEEMDAN denoising model (Figure 1) consists of the following steps:

(i) First, the nodal seismograph signal is decomposed using ICEEMDAN to obtain several IMF components. It is necessary to further distinguish between effective and noisy components before performing IMF reconstruction.

(ii) The COEF, MI, $R^2$, Adj-$R^2$, JSD, CS, RMSE, MAE, MAPE, and sample entropy for each of the above IMF components are calculated to construct a 10-dimensional IMF evaluation index matrix.

(iii) GRA is performed on the IMF evaluation index matrix, followed by a sequential standardization, calculation of the minimum/maximum difference between two levels, calculation of grey relational coefficients, index weight estimation, and IMF correlation analysis to obtain the final ranking of all IMF components.

(iv) Finally, the top half of the ranked IMF components is linearly constructed to obtain the denoised nodal seismograph data.

To test the feasibility and denoising effect of the above GRA–ICEEMDAN process, it is necessary to combine some simulation experiments for evaluation. The specific details will be discussed in the experimental analysis section of Sections 3 and 4.

## 3. Experiment and analysis of random noise suppression for simulated noisy signals

To test the theoretical denoising performance of the GRA–ICEEMDAN model, a simulated noisy signal denoising experiment was introduced. Ten metrics—COEF, sample entropy, MI, $R^2$, CS, JSD, RMSE, MAE, Adj-$R^2$, and MAPE—were used to evaluate the quality of several IMF components obtained from the ICEEMDAN decomposition. Then, the denoising model's effectiveness was evaluated using three metrics: COEF, SNR, and MSE. Finally, the denoising performance was compared with several classic models.

The simulation platform used was MATLAB 2024b, with a signal length of 10 s and a sampling rate of 500 Hz. The simulated seismic signal consists of three waveforms (fundamental wave, P-wave, S-wave). The fundamental

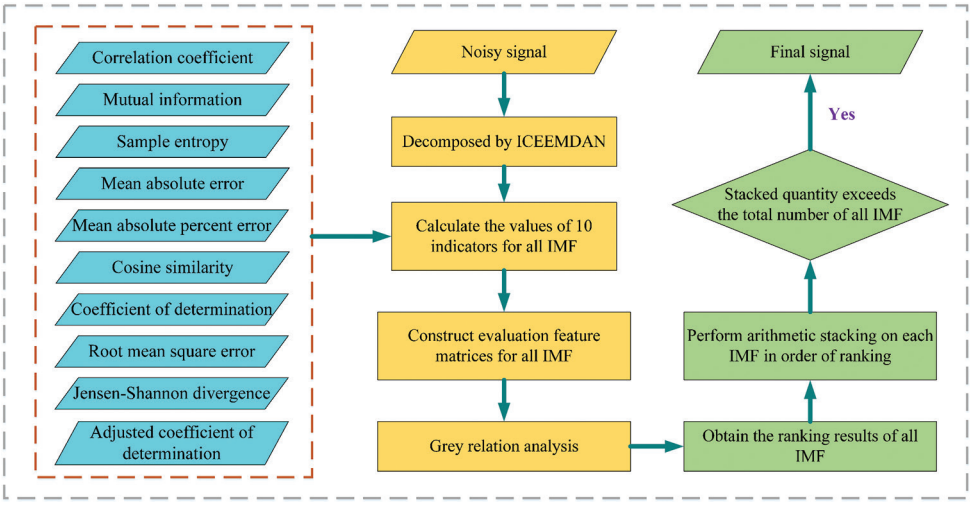



**Figure 1.** Denoising process of the grey relational analysis–improved complete ensemble empirical mode decomposition with adaptive noise signal denoising model
Abbreviation: IMF: Intrinsic mode function.

wave simulates background microseismic components (low-frequency continuous signal). It generates the base waveform by superimposing multiple low-frequency sine waves across the entire time domain, with the generation function defined as **Equation 18**:

$$y(t) = 1 + \sum_{i=1}^{3} \left( \alpha_i \times \sin\left( 2\pi \cdot f_i^{back} \cdot t \right) \right) \tag{18}$$

where $\alpha_i$ = 0.3 or 0.7 or 1.2, and, $f_i^{back} = 0.02 \ or \ 0.06 \ or \ 0.01$.

The P-wave is designed to start at 1 s and last for 1.8 s. It is formed by the Hadamard product of a power function and a sine function, simulating a high-frequency waveform with short duration and relatively small amplitude, that is, $y_p(t_p) = 0.6 \times e^{\alpha_p t_p} \cdot \sin\left( 2\pi \cdot f_p \cdot t_p \right)$. The S-wave was designed to start at 2.5 s and last for 4.8 s, composed of the Hadamard product of a power function and a sine function, simulating a low-frequency waveform with larger amplitude and longer duration, that is, $y_s(t_s) = 2 \times (e^{\alpha_s t_s} \cdot (1 - e^{\alpha_s t_s})) \cdot \sin\left( 2\pi \cdot f_s \cdot t_s \right)$. By adding 20 SNR/dB of Gaussian white noise to the simulated seismic signal using MATLAB's built-in "awgn" function, the noise-containing signal required for the experiment was generated.

The ICEEMDAN decomposition parameters were set as follows: Noise intensity 0.2, number of integrations 3, and maximum iterations 3,600. The similarity tolerance threshold for sample entropy was defined as 0.15 times the signal standard deviation. The resulting clean signal, noisy signal, and corresponding ICEEMDAN decomposition plots are shown in the subplots of Figure 2.

Table 1 presents the eigenvalues of each IMF component after decomposing the original simulated noisy signal using ICEEMDAN, encompassing 10 distinct metric values. $IMF_1$ to $IMF_2$, obtained from decomposing the noisy signal, exhibited high-frequency, low-amplitude random oscillations. Their sample entropy values exceeded 1.5, and their COEFs with the original signal were all below 0.1, with the various error function values (RMSE, MAE, etc.) being relatively large. These characteristics indicate significant amplitude differences between these IMFs and the original signal, suggesting that they represent system noise components that should be removed. $IMF_3$–$IMF_4$ exhibited smoother waveforms (sample entropies below 0.06) with pronounced low-frequency characteristics. Their amplitude magnitudes and curve shapes (COEFs exceeding 0.5) aligned more closely with the original signal, qualifying them as valid components extracted from the noisy original signal. The amplitudes of other IMFs were all below 0.1, showing a significant gap compared to the peak value of the original signal (1.9). Whether they need to be retained requires correlation analysis using GRA.

The IMF quality evaluation indicator matrix was derived from Table 1. After undergoing necessary processing steps—including standardization, two-stage least squares/maximum differences calculation, grey relational coefficient computation, subjective assignment of indicator weights, and IMF correlation analysis—the final quality ranking results for IMF components were obtained. The specific calculation results are shown in Tables 2 and 3 and Figure 3. Table 2 presents the comparative sequences and reference sequence values for each IMF in the simulated noisy signal, while Table 3 displays the GRA COEFs and

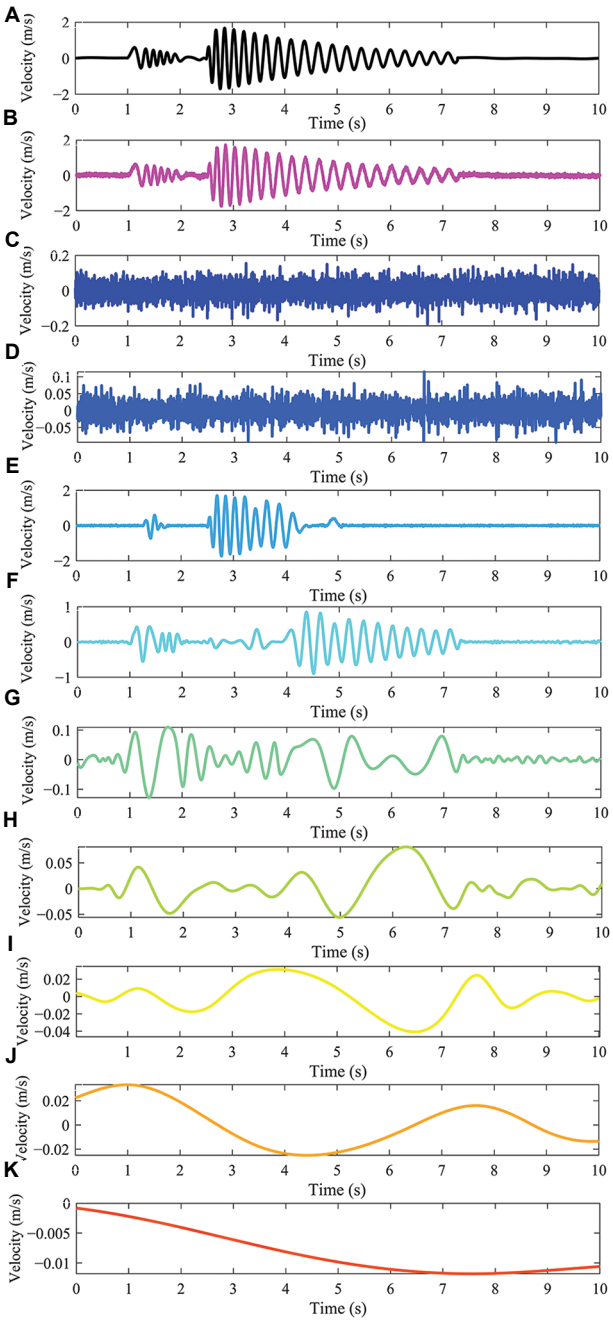



**Figure 2.** Simulated original signal and improved complete ensemble empirical mode decomposition with adaptive noise signal denoising decomposition results. (A) Original clean signal. (B) Original noisy signal. (C) $IMF_1$. (D) $IMF_2$. (E) $IMF_3$. (F) $IMF_4$. (G) $IMF_5$. (H) $IMF_6$. (I) $IMF_7$. (J) $IMF_8$. (K) $IMF_9$.
Abbreviation: IMF: Intrinsic mode function.

correlation values for each IMF in the simulated noisy signal.

The comparison and reference sequences in Table 2 can be used to construct the GRA COEFs between 10 evaluation indicators and each IMF (Table 3). Based on the balanced indicator weight vector {0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1} and the GRA COEFs, the grey weighted correlation degree of each IMF was obtained using arithmetic average (Table 3).

$IMF_3$, $IMF_8$, and $IMF_9$ achieved the best performance (COEF value of 1) in seven, one, and two out of 10 evaluation metrics, respectively, indicating that $IMF_8$ exhibited the most significant correlation with the original signal. Specifically, $IMF_3$ attained the highest scores in seven metrics, including COEF, Adj-$R^2$, MAE, CS, RMSE, $R^2$, and JSD correlation; $IMF_8$ only outperformed in the MI metric, while $IMF_9$ stood out in metrics, such as sample entropy and MAPE. In the comprehensive ranking ($IMF_3$ > $IMF_4$ > $IMF_9$ > $IMF_7$ > $IMF_8$ > $IMF_5$ > $IMF_6$ > $IMF_2$ > $IMF_1$), although $IMF_4$ did not achieve the optimal GRA COEF in any single metric, its COEFs exceeded 0.5 across six metrics, indicating it also captured key signal components. Therefore, the top five IMF models ($IMF_3$, $IMF_4$, $IMF_7$, $IMF_8$, $IMF_9$) were selected for linear reconstruction.

To compare the GRA–ICEEMDAN model's superiority among similar signal denoising principles, nine reference models—wavelet threshold denoising, median filter, Wiener filter, Savitzky–Golay filter, elliptic filter, Butterworth filter, Kalman filter, Chebyshev filter, and moving average filter—were used. The evaluation framework incorporated quantitative metrics across multiple dimensions, including SNR, noise suppression ratio, correlation, frequency similarity, energy retention, and overall score. The denoising results for each model are presented in Figure 3.

Compared to the significant phase delays and severe waveform distortion observed in other models, the wavelet, median filter, Wiener filter, and GRA–ICEEMDAN models demonstrated positive effects in seismic signal denoising, achieving a degree of smoothing and high time-frequency fidelity (Figure 3A-C). The 0.5-s local magnification (Figure 3C) reveals substantial differences in local characteristics among the 10 denoising models, with varying degrees of divergence in amplitude magnitude, phase delay, frequency, and peak energy. Among these, the wavelet and GRA–ICEEMDAN models exhibited smoother and more stable waveform processing at peaks or troughs, while the Chebyshev denoising results completely lost all original signal characteristics. Issues with other models were primarily centered on phase delays (0.025–0.05 s) and amplitude reductions.

This scale disparity was also pronounced in the spectral plots, with spectral curve differences among models primarily concentrated in the low-to-mid frequency

**Table 1. Characteristic index values of each IMF after ICEEMDAN decomposition of the simulated noisy signal**

| IMF components | Correlation coefficient | Sample entropy | CS | $R^2$ | JSD | RMSE | MAE | MAPE | Adj-$R^2$ | MI |
|---|---|---|---|---|---|---|---|---|---|---|
| $IMF_1$ | 0.0719 | 2.2866 | 0.0718 | 0.0047 | 0.4131 | 0.4682 | 0.2924 | 2.8243 | 0.0049 | 0.8379 |
| $IMF_2$ | 0.0437 | 1.5415 | 0.0439 | 0.0017 | 0.4242 | 0.4689 | 0.2961 | 2.0493 | 0.0017 | 0.8355 |
| $IMF_3$ | 0.8479 | 0.0198 | 0.8479 | 0.7189 | 0.1692 | 0.2488 | 0.1665 | 6.5080 | 0.7189 | 0.6970 |
| $IMF_4$ | 0.5247 | 0.0544 | 0.5248 | 0.2750 | 0.2080 | 0.3996 | 0.1921 | 8.1701 | 0.2752 | 0.7920 |
| $IMF_5$ | 0.0210 | 0.0762 | 0.0216 | −0.0039 | 0.3815 | 0.4702 | 0.2958 | 2.7874 | 0.0002 | 0.8423 |
| $IMF_6$ | 0.0084 | 0.0266 | 0.0094 | −0.0027 | 0.4372 | 0.4699 | 0.2972 | 2.8712 | −0.0001 | 0.8473 |
| $IMF_7$ | 0.0190 | 0.0142 | 0.0186 | −0.0001 | 0.3512 | 0.4693 | 0.2968 | 2.1045 | 0.0001 | 0.8655 |
| $IMF_8$ | 0.0282 | 0.0049 | 0.0288 | 0.0006 | 0.5020 | 0.4691 | 0.2964 | 2.4613 | 0.0005 | 0.8714 |
| $IMF_9$ | 0.0283 | 0.0007 | 0.0025 | −0.0003 | 0.3233 | 0.4694 | 0.2974 | 1.7733 | 0.0006 | 0.8653 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

**Table 2. Comparison series and reference series values of each IMF of the simulated noisy signal**

| Serial no. | Evaluation indicators | $IMF_1$ | $IMF_2$ | $IMF_3$ | $IMF_4$ | $IMF_5$ | $IMF_6$ | $IMF_7$ | $IMF_8$ | $IMF_9$ | Optimal IMF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Correlation coefficient | 0.0756 | 0.0419 | 1 | 0.6149 | 0.0149 | 0 | 0.0125 | 0.0235 | 0.0237 | 1 |
| 2 | Sample entropy | 0 | 0.3259 | 0.9916 | 0.9765 | 0.9670 | 0.9886 | 0.9940 | 0.9981 | 1 | 1 |
| 3 | CS | 0.0819 | 0.0489 | 1 | 0.6177 | 0.0226 | 0.0083 | 0.0190 | 0.0311 | 0 | 1 |
| 4 | $R^2$ | 0.0121 | 0.0079 | 1 | 0.3859 | 0 | 0.0016 | 0.0052 | 0.0063 | 0.0049 | 1 |
| 5 | JSD | 0.2671 | 0.2338 | 1 | 0.8833 | 0.3620 | 0.1945 | 0.4531 | 0 | 0.5368 | 1 |
| 6 | RMSE | 0.0092 | 0.006088 | 1 | 0.3190 | 0 | 0.0012 | 0.0039 | 0.0048 | 0.0038 | 1 |
| 7 | MAE | 0.0378 | 0.0097 | 1 | 0.8041 | 0.0118 | 0.0010 | 0.0040 | 0.0075 | 0 | 1 |
| 8 | MAPE | 0.8356 | 0.9568 | 0.2598 | 0 | 0.8414 | 0.8283 | 0.9482 | 0.8924 | 1 | 1 |
| 9 | Adj-$R^2$ | 0.0071 | 0.0025 | 1 | 0.3828 | 0.0005 | 0 | 0.0004 | 0.0010 | 0.0010 | 1 |
| 10 | MI | 0.8077 | 0.7938 | 0 | 0.5446 | 0.8330 | 0.8618 | 0.9663 | 1 | 0.9651 | 1 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

**Table 3. Grey relation analysis, correlation coefficients, and correlation values of each IMF in the simulated noisy signal**

| Serial number | Evaluation indicators | The IMF evaluated | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $IMF_1$ | $IMF_2$ | $IMF_3$ | $IMF_4$ | $IMF_5$ | $IMF_6$ | $IMF_7$ | $IMF_8$ | $IMF_9$ |
| 1 | Correlation coefficient | 0.3510 | 0.3429 | 1 | 0.5649 | 0.3366 | 0.3333 | 0.3361 | 0.3386 | 0.3386 |
| 2 | Sample entropy | 0.3333 | 0.4258 | 0.9836 | 0.9551 | 0.9380 | 0.9778 | 0.9883 | 0.9963 | 1 |
| 3 | CS | 0.3526 | 0.3445 | 1 | 0.5667 | 0.3384 | 0.3351 | 0.3376 | 0.3404 | 0.3333 |
| 4 | $R^2$ | 0.3360 | 0.3351 | 1 | 0.4488 | 0.3333 | 0.3336 | 0.3344 | 0.3347 | 0.3344 |
| 5 | JSD | 0.4055 | 0.3948 | 1 | 0.8108 | 0.4393 | 0.3829 | 0.4776 | 0.3333 | 0.5191 |
| 6 | RMSE | 0.3354 | 0.3346 | 1 | 0.4233 | 0.3333 | 0.3336 | 0.3342 | 0.3344 | 0.3341 |
| 7 | MAE | 0.3419 | 0.3355 | 1 | 0.7185 | 0.3359 | 0.3335 | 0.3342 | 0.3350 | 0.3333 |
| 8 | MAPE | 0.7526 | 0.9205 | 0.4031 | 0.3333 | 0.7592 | 0.7444 | 0.9061 | 0.8229 | 1 |
| 9 | Adj-$R^2$ | 0.3349 | 0.3339 | 1 | 0.4475 | 0.3334 | 0.3333 | 0.3334 | 0.3335 | 0.3335 |
| 10 | MI | 0.7223 | 0.7080 | 0.3333 | 0.5233 | 0.7496 | 0.7835 | 0.9370 | 1 | 0.9348 |
| 11 | Correlation | 0.4265 | 0.4476 | 0.8720 | 0.5792 | 0.4897 | 0.4891 | 0.5319 | 0.5169 | 0.5461 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

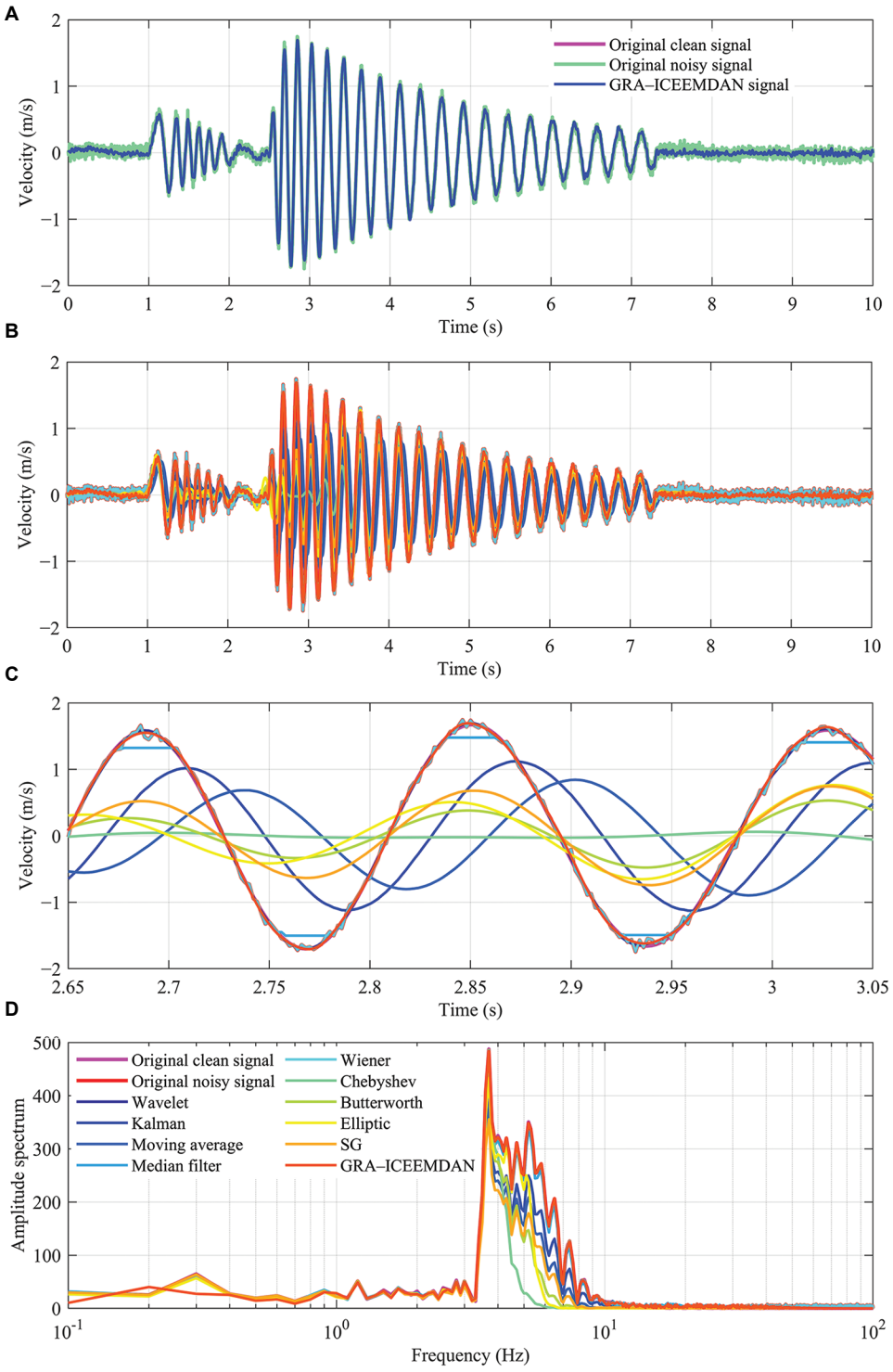



**Figure 3.** Comparison of waveforms before and after noise reduction for simulated seismic signals under different noise reduction models. Comparison of (A) denoising results between the original noisy signal and GRA–ICEEMDAN, (B) results from multiple noise reduction methods, (C) localized amplified waveforms from multiple noise reduction models, and (D) results from multiple denoising models (0.1–100 Hz).
Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SG: Savitzky–Golay.

ranges of 0.1–0.4 Hz and 3.5–10 Hz. Spectral consistency in other regions generally remained at an acceptable level. Therefore, when processing complex waveforms, such as simulated seismic waves, most models produced noise rejection misjudgments, failing to accurately locate the phase position and main peak scale of the effective signal. This leads to significant misjudgments in both the time domain and frequency domain. The excellent performance of the wavelet and GRA–ICEEMDAN models on simulated waveforms offers new insights and application prospects for seismic data processing. Further evaluation of their applicability scope requires incorporating comparative analysis with real seismic signals.

Figure 4 presents bar charts comparing denoising metrics for simulated seismic signals across different denoising models. This visualization clearly reveals the performance differences among models in both overall metrics and specific indicators. The combined score bars for the GRA–ICEEMDAN and wavelet models are tied at the highest position, clearly demonstrating their leading advantage in overall denoising performance. Further examination of the sub-metric bar charts revealed that despite their identical total scores, their performance characteristics diverged. The GRA–ICEEMDAN model maintained high levels across SNR, correlation, frequency similarity, and energy retention, forming a relatively balanced "performance spectrum". However, its bar height for noise suppression rate is significantly lower than that of other high-scoring models. This apparent "weakness" aligns with its algorithm's restrained approach to noise removal—prioritizing signal fidelity over aggressive noise elimination to prevent excessive distortion.

In contrast, the composite score bar for the moving average and Chebyshev models clearly ranked in the lowest tier. Their sub-metric bar charts exhibited a distinctly polarized pattern: The bar for noise suppression rate was exceptionally prominent, nearly reaching full value, while the bars for SNR and energy retention rate were severely depressed. This stark contrast in the bar charts directly visualizes the flaw of "excessive denoising" severely damaging the effective signal, explaining why these models rank at the bottom. The bar shapes for Butterworth, elliptic, Savitzky–Golay, Wiener, and Kalman models also exhibited similar imbalances, albeit to varying degrees. These bar charts clearly revealed the distribution characteristics of performance differences among models and their strengths and weaknesses across different dimensions, enhancing the objectivity and persuasiveness of the discussion on denoising model results. Detailed denoising evaluation metric data are fully presented in Table 4.

Based on the overall score ranking results (GRA–ICEEMDAN=Wavelet>Median filter>Wiener>Savitzky–Golay>Elliptic filter>Butterworth filter>Kalman filter>Chebyshev filter>Moving average), it can be observed that the GRA–ICEEMDAN model and wavelet denoising model demonstrated superior overall noise suppression performance, both achieving a perfect score of 100 (out of 100). Their performances across individual evaluation metrics were also relatively balanced. Specifically, the GRA–ICEEMDAN model's sub-metric scores were 24.0049 dB for SNR, 87.9105% for noise suppression rate, 0.9946 for signal correlation, 99.7962% for frequency similarity, and 99.0641% for energy retention. Compared to the white noise power introduced into the clean signal (20 dB), it achieved an improvement of 4.0049 dB. The entire denoising process took 3.7754 s to compute. Each metric contributed equally to the overall calculation. Among the 10 denoising models, it ranked second, eighth, fourth, and second in SNR, noise

**Table 4. Comparison of the denoising effects of the models based on simulated denoised data**

| Denoising model | SNR (db) | Noise suppression ratio (%) | Correlation | Frequency similarity (%) | Energy retention (%) | Overall score |
|---|---|---|---|---|---|---|
| Wavelet | 30.5713 | 97.2009 | 0.9952 | 99.8684 | 98.9768 | 100 |
| Kalman filter | 3.8652 | 99.3746 | 0.7650 | 99.4674 | 55.6142 | 71.7661 |
| Moving average | −1.0326 | 99.8781 | 0.1233 | 97.7523 | 42.6740 | 52.2340 |
| Median filter | 23.3959 | 81.6674 | 0.9948 | 99.8196 | 92.3022 | 98.1086 |
| Wiener filter | 20.0868 | 2.0009 | 0.9999 | 99.9999 | 99.8454 | 80.4561 |
| Chebyshev filter | 3.0365 | 99.9982 | 0.7163 | 71.7636 | 34.6873 | 62.1236 |
| Butterworth filter | 5.8201 | 99.9998 | 0.8871 | 89.0363 | 42.0791 | 72.5777 |
| Elliptic filter | 6.3841 | 99.9998 | 0.8820 | 88.5631 | 57.0188 | 74.7498 |
| Savitzky–Golay | 6.7753 | 99.9045 | 0.9591 | 96.3679 | 34.4964 | 75.6013 |
| GRA–ICEEMDAN | 24.0049 | 87.9105 | 0.9946 | 99.7962 | 99.0641 | 100 |

Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SNR: Signal-to-noise ratio.

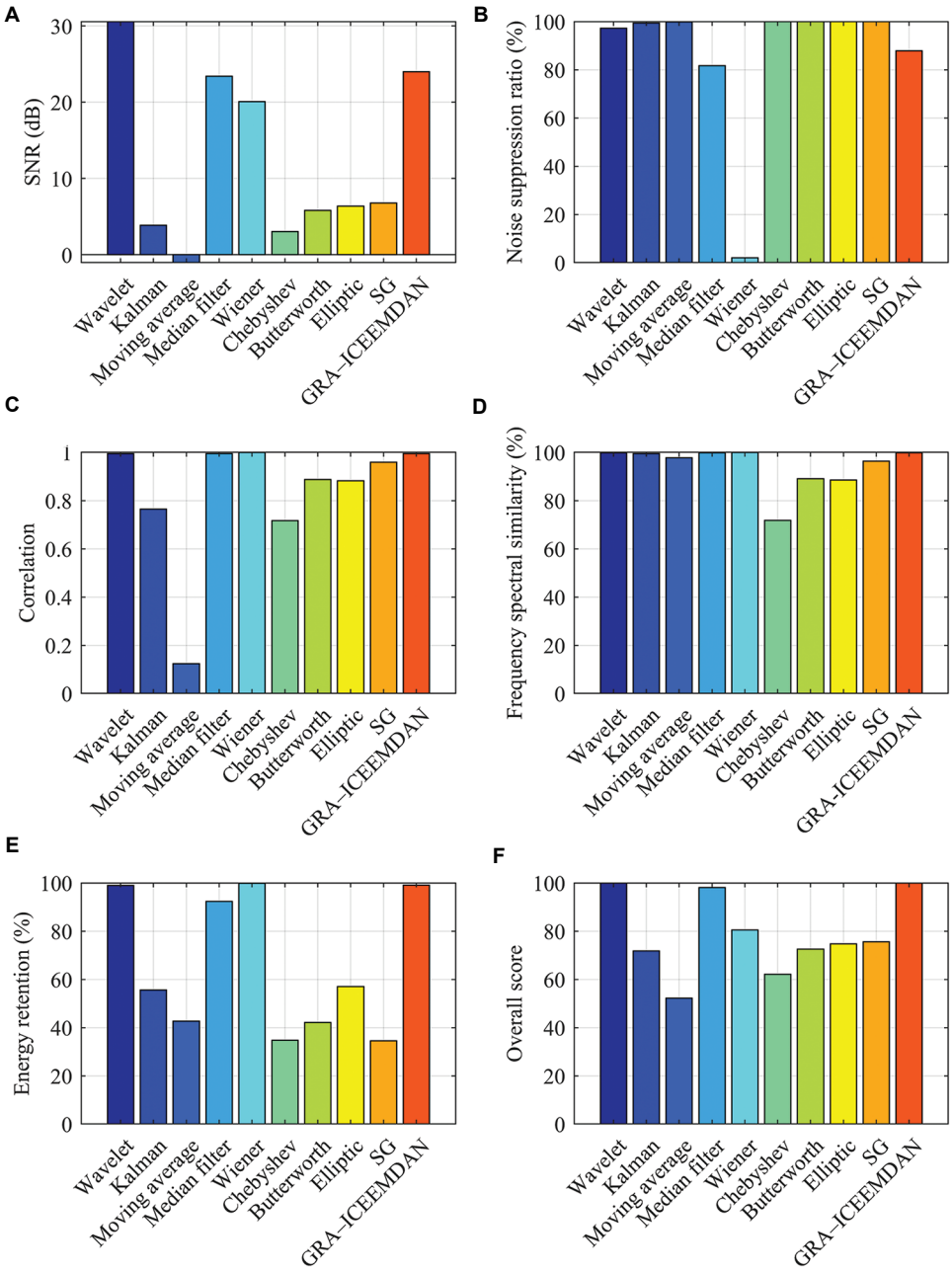



**Figure 4.** Bar chart comparison of noise reduction metrics for simulated seismic signals under different noise reduction models. (A) Signal-to-noise ratio (SNR) improvement. (B) Noise suppression ratio. (C) Correlation. (D) Frequency spectral similarity. (E) Energy retention. (F) Overall score.
Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SG: Savitzky–Golay.

suppression rate, signal correlation, frequency spectrum similarity, and energy retention, respectively, with most individual metrics ranking highly. However, its ranking for noise suppression ratio (87.9105%) was low. This is because the GRA–ICEEMDAN model automatically balances the relationship between main frequency energy retention, waveform consistency, and noise suppression. It controls noise removal within manageable limits to prevent energy leakage or severe amplitude distortion.

Some models exhibited poor overall noise reduction performance, such as the moving average and Chebyshev models, which demonstrate subpar denoising capabilities (overall score < 65). Their post-denoising performance and waveform characteristics differed significantly from

those of other models. Furthermore, their performance metrics (noise suppression ratio of 99.8781% and 99.9982%, respectively) confirmed that these models suffer from excessive denoising, rendering them incapable of accurately assessing noise suppression levels. (SNR < 4 dB, Energy retention < 45%). Consequently, they lack competitiveness in overall score rankings. Models, such as Butterworth, elliptic, Savitzky–Golay, Wiener, and Kalman exhibited similar noise reduction flaws, all displaying varying degrees of excessive denoising. This directly leads to severe distortion of the original waveform and phase confusion, imposing certain limitations on their application in reliable seismic monitoring and geophysical exploration projects.

## 4. Experiment and analysis of nodal seismograph signal denoising

The IGU-16HR nodal seismograph is a highly representative wireless and velocity-type intelligent instrument in the fields of geophysical exploration and seismic monitoring. It possesses advantages such as low cost, long battery life, small size, and rapid deployment. It can be equipped with various tools, including a mobile deployment terminal, a smart app, a charging box, a data download box, and data pre-processing desktop software, to enhance its field observation capabilities. The main performance parameters of the 5 Hz instrument are support for continuous data acquisition for 35 days, distortion less than 0.1%, spurious frequency greater than 170 Hz, open-circuit intrinsic voltage sensitivity of 80 V/m/s, sampling interval support of 0.25–4 ms, IP68 waterproof rating, built-in 206.5 Hz anti-aliasing filter and 1–10 Hz DC blocking filter, system dynamic range of 145 dB, and frequency response of 0–1,652 Hz. Given the superior performance of the IGU-16HR nodal seismograph, this study mainly utilizes data recorded by this model to conduct various denoising experiments and analyses.

The GRA–ICEEMDAN denoising experimental data mainly came from the Ms 3.2 earthquake event recorded on June 27, 2025, in Lichuan, Hubei, China, by a single-component SmartSolo IGU-16HR nodal seismograph, and the geopulsation environmental excitation event recorded on May 31, 2025, at a seismic station on the campus of Wuhan University in Wuhan, Hubei. The sampling rates were 500 Hz and 1,000 Hz, respectively. The former signal successfully recorded the S-coda wave of the earthquake event, and the latter signal successfully recorded the regular geopulsation information of the city at night. Experimental platform and algorithm parameters: The simulation platform used was MATLAB 2024b, and the signal length was 10 s. The noise intensity, number of ensembles, and maximum number of iterations for ICEEMDAN decomposition were 0.2, 3, and 3,600, respectively. The similarity tolerance threshold for sample entropy was the product of 0.15 and the signal standard deviation.

### 4.1. Denoising experiment of S-coda waves recorded by nodal seismograph

First, GRA–ICEEMDAN denoising analysis was performed on the S-coda waves of natural seismic events recorded by node seismometers. The signal sampling rate was 500 Hz, and the experimental data duration was 10 s. The GRA analysis process data are shown in Tables 5-7, the signal decomposition results are shown in Figure 5, and the denoising results under various models are shown in Table 8 and Figures 6-8.

Figure 5 shows the schematic diagram of the eight IMF components obtained by decomposing the seismic signal recorded by the node seismometer using ICEEMDAN. From Figure 5, it can be observed that $IMF_1$ to $IMF_4$ contained substantial noise with overall minimal amplitudes, exhibiting generally high oscillation frequencies around the zero-time point throughout the entire time domain.

**Table 5. Characteristic parameter values of each IMF in the S-coda waves of seismic events recorded by nodal seismographs**

| IMF | Correlation coefficient | Sample entropy | CS | $R^2$ | JSD | RMSE | MAE | MAPE | Adj-$R^2$ | MI |
|---|---|---|---|---|---|---|---|---|---|---|
| $IMF_1$ | 0.3213 | 0.3936 | 0.1564 | −3.1168 | 0.3256 | 1.3282e−07 | 1.1774e−07 | 1.2357 | 0.1030 | 0.7722 |
| $IMF_2$ | 0.3111 | 0.9169 | 0.1541 | −3.1244 | 0.2510 | 1.3294e−07 | 1.1777e−07 | 1.1729 | 0.0966 | 0.8721 |
| $IMF_3$ | 0.0921 | 0.6396 | 0.0420 | −3.2106 | 0.2842 | 1.3433e−07 | 1.1825e−07 | 1.0660 | 0.0082 | 0.8558 |
| $IMF_4$ | 0.0100 | 0.5480 | 0.0132 | −3.2174 | 0.2931 | 1.3444e−07 | 1.1829e−07 | 1.0093 | −9.9532e−05 | 0.8784 |
| $IMF_5$ | 0.0766 | 0.1629 | 0.0338 | −3.2137 | 0.2787 | 1.3438e−07 | 1.1835e−07 | 1.0232 | 0.0056 | 0.8792 |
| $IMF_6$ | 0.4057 | 0.0372 | 0.0083 | −3.2357 | 0.2753 | 1.3473e−07 | 1.1997e−07 | 1.3947 | 0.1644 | 0.8835 |
| $IMF_7$ | 0.5731 | 0.0138 | 0.2251 | −3.0392 | 0.1844 | 1.3157e−07 | 1.2082e−07 | 1.8793 | 0.3284 | 0.9049 |
| $IMF_8$ | 0.5723 | 0.0031 | 0.9160 | 0.3101 | 0.0348 | 5.4372e−08 | 4.6263e−08 | 1.9774 | 0.3274 | 0.9099 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

**Table 6. Comparison and reference series values of each IMF for S-coda waves of seismic events recorded by nodal seismographs**

| Index | The IMF evaluated | | | | | | | | Ideal IMF |
|---|---|---|---|---|---|---|---|---|---|
| | $IMF_1$ | $IMF_2$ | $IMF_3$ | $IMF_4$ | $IMF_5$ | $IMF_6$ | $IMF_7$ | $IMF_8$ | |
| Correlation coefficient | 0.5527 | 0.5346 | 0.14579 | 0 | 0.1182 | 0.7026 | 1 | 0.9984 | 1 |
| Sample entropy | 0.5727 | 0 | 0.3034 | 0.4036 | 0.8250 | 0.9626 | 0.9882 | 1 | 1 |
| CS | 0.1631 | 0.1606 | 0.0371 | 0.0053 | 0.0280 | 0 | 0.2388 | 1 | 1 |
| $R^2$ | 0.0335 | 0.0313 | 0.0070 | 0.0051 | 0.0062 | 0 | 0.0554 | 1 | 1 |
| JSD | 0 | 0.2565 | 0.1424 | 0.1118 | 0.1613 | 0.1731 | 0.4855 | 1 | 1 |
| RMSE | 0.0237 | 0.0221 | 0.0049 | 0.0036 | 0.0043 | 0 | 0.0393 | 1 | 1 |
| MAE | 0.0413 | 0.0408 | 0.0344 | 0.0339 | 0.0332 | 0.0114 | 0 | 1 | 1 |
| MAPE | 0.7661 | 0.8310 | 0.9414 | 1 | 0.9856 | 0.6018 | 0.1013 | 0 | 1 |
| Adj-$R^2$ | 0.3140 | 0.2944 | 0.0255 | 0 | 0.0175 | 0.5008 | 1 | 0.9969 | 1 |
| MI | 0 | 0.7251 | 0.6073 | 0.7708 | 0.7768 | 0.8085 | 0.9638 | 1 | 1 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

**Table 7. Grey relational analysis correlation coefficients and correlation values of each IMF in the S-coda waves of seismic events recorded by nodal seismographs**

| Evaluation indicators | The IMF evaluated | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $IMF_1$ | $IMF_2$ | $IMF_3$ | $IMF_4$ | $IMF_5$ | $IMF_6$ | $IMF_7$ | $IMF_8$ |
| Correlation coefficient | 0.5278 | 0.5179 | 0.3692 | 0.3333 | 0.3618 | 0.6270 | 1 | 0.9969 |
| Sample entropy | 0.5391 | 0.3333 | 0.4178 | 0.4560 | 0.7408 | 0.9304 | 0.9769 | 1 |
| CS | 0.3740 | 0.3733 | 0.3418 | 0.3345 | 0.3396 | 0.3333 | 0.3964 | 1 |
| $R^2$ | 0.3409 | 0.3404 | 0.3349 | 0.3344 | 0.3347 | 0.3333 | 0.3461 | 1 |
| JSD | 0.3333 | 0.4021 | 0.3683 | 0.3601 | 0.3735 | 0.3768 | 0.4928 | 1 |
| RMSE | 0.3387 | 0.3383 | 0.3344 | 0.3341 | 0.3343 | 0.3333 | 0.3423 | 1 |
| MAE | 0.3427 | 0.3426 | 0.3411 | 0.3410 | 0.3408 | 0.3358 | 0.3333 | 1 |
| MAPE | 0.6813 | 0.7474 | 0.8951 | 1 | 0.9720 | 0.5567 | 0.3574 | 0.3333 |
| Adj-$R^2$ | 0.4215 | 0.4147 | 0.3391 | 0.3333 | 0.3372 | 0.5004 | 1 | 0.9940 |
| MI | 0.3333 | 0.6452 | 0.5601 | 0.6857 | 0.6914 | 0.7230 | 0.9325 | 1 |
| Correlation | 0.4233 | 0.4455 | 0.4302 | 0.4512 | 0.4826 | 0.5050 | 0.6178 | 0.9324 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

The curves of $IMF_5$ to $IMF_8$ demonstrated significantly smoother and more stable characteristics compared to $IMF_1$ to $IMF_3$, showing no concentration of energy, with amplitudes similarly remaining within a low range. $IMF_7$ exhibited high similarity to the original waveform, a similarity supported by objective assessment through COEFs or mutual information metrics. After obtaining the ICEEMDAN decomposition results for seismic event signals recorded by node seismometers (Figure 5), 10 characteristic index values were calculated for each IMF, as presented in Table 5.

Table 5 presents the evaluation metrics for each IMF component following the ICEEMDAN decomposition of node seismometer signals. It is evident that $IMF_6$ to $IMF_8$ exhibited significant differences from other IMF components across multiple evaluation metrics, such as sample entropy, Adj-$R^2$, COEF, and mutual information. Particularly, numerous eigenvalues of the $IMF_8$ component exhibited extreme contrasts in magnitude compared to other IMFs, while simultaneously presenting minimal values in metrics, such as RMSE and MAE, which express the absolute error between two signals.

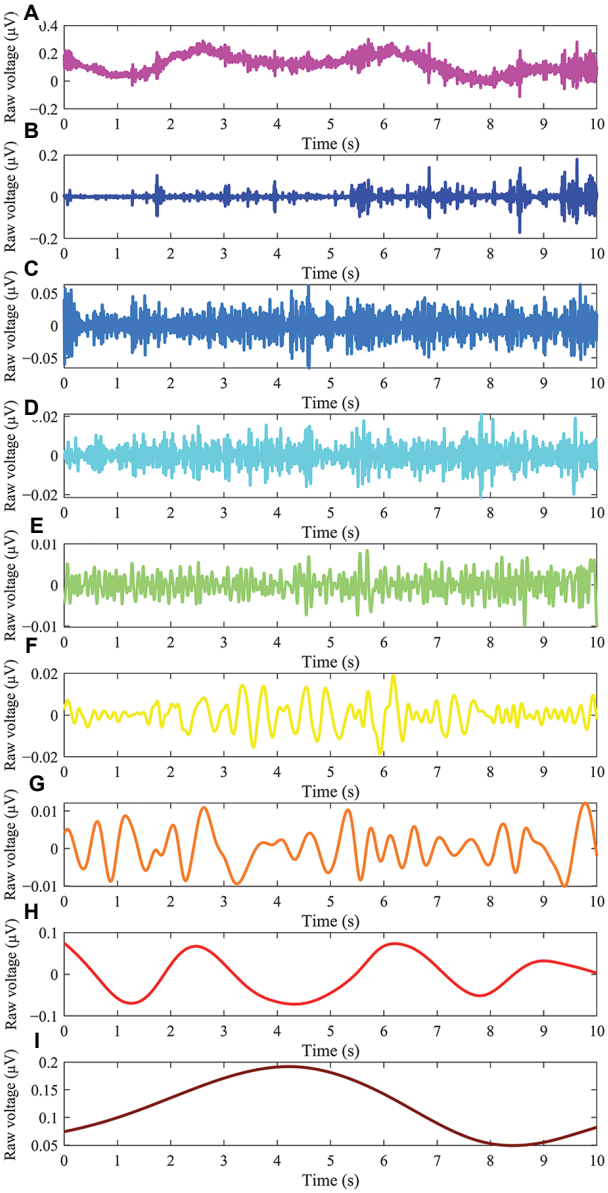



**Figure 5.** Schematic diagram of eight IMFs obtained by decomposing the S-coda wave of a seismic event recorded by a nodal seismograph using ICEEMDAN. (A) Original noisy signal. (B) $IMF_1$. (C) $IMF_2$. (D) $IMF_3$. (E) $IMF_4$. (F) $IMF_5$. (G) $IMF_6$. (H) $IMF_7$. (I) $IMF_8$.
Abbreviation: ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; IMF: Intrinsic mode function.

The IMF quality evaluation metric matrix is derived from Table 5. After undergoing standardization, two-stage least squares/two-stage maximum difference calculation, grey relational coefficient calculation, subjective assignment of indicator weights, and IMF correlation analysis, the final quality ranking results for IMF components are presented in Tables 6 and 7. Table 6 shows the comparison sequences and reference sequence values for each IMF of the seismic event signal recorded by the node seismometer, while Table 7 presents the GRA COEFs and correlation values for each IMF of the seismic event signal recorded by the node seismometer.

The comparison and reference sequences in Table 6 can be used to construct the GRA COEFs between 10 evaluation indicators and each IMF (Table 7). Based on the balanced indicator weight vector {0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1} and the GRA COEFs, the grey weighted correlation degree of each IMF was obtained using the arithmetic average (Table 7).

The final GRA correlation ranking results are: $IMF_8$ > $IMF_7$ > $IMF_6$ > $IMF_5$ > $IMF_4$ > $IMF_2$ > $IMF_3$ > $IMF_1$, with GRA scores of 0.9324, 0.6178, 0.5050, 0.4826, 0.4512, 0.4455, 0.4302, and 0.4233, respectively. As shown in Table 7, among all intrinsic modal functions, $IMF_8$ exhibited the highest correlation (0.9324) with the ideal IMF and possessed the optimal COEF (value of 1) for seven indicators, proving its superiority over any other IMF. $IMF_7$ follows with two optimal COEFs, and $IMF_4$ with one optimal COEF. Therefore, the top four IMF rankings ($IMF_8$, $IMF_7$, $IMF_6$, $IMF_5$) were selected for linear reconstruction, with results shown in Figure 6.

To compare the GRA–ICEEMDAN model's superiority among similar signal denoising models, nine reference models—wavelet threshold denoising, median filter, Wiener filter, Savitzky–Golay filter, elliptic filter, Butterworth filter, Kalman filter, Chebyshev filter, and moving average filter— were used. The evaluation framework incorporated quantitative metrics across multiple dimensions, including SNR improvement, noise suppression ratio, correlation, frequency similarity, energy retention, and overall score. The denoising results for each model are also displayed in Figure 6.

Except for the wavelet and Wiener models, other models produced positive effects in seismic signal denoising, achieving a certain degree of smoothing and high time-frequency fidelity (Figure 6A, B, and D), though they exhibited notable differences in waveform processing. Due to the presence of strong and widely distributed high-frequency noise in natural seismic signals, the 0.5-s local magnification view (Figure 6C) revealed that Kalman filtering, median filtering, and the moving average model retained some characteristics of high-frequency noise, causing severe phase distortion. In contrast, models such as GRA–ICEEMDAN, elliptic, Savitzky–Golay, and Butterworth models demonstrated superior high-frequency noise suppression. Their local waveforms exhibited smoother, more stable trajectories aligned with the original signal's progression and clustered near the center of the high-frequency signal curve.

**Table 8. Comparison of nodal seismograph signal denoising effects under different models**

| Denoising model | SNR improvement (db) | Noise suppression ratio (%) | Correlation | Frequency similarity (%) | Energy retention (%) | Overall score |
|---|---|---|---|---|---|---|
| Wavelet | 0.1605 | 3.6695 | 0.9991 | 99.9874 | 99.8499 | 60.8449 |
| Kalman filter | 11.9065 | 94.2601 | 0.9284 | 99.0835 | 96.3058 | 88.4194 |
| Moving average | 19.1154 | 99.4890 | 0.9053 | 98.6071 | 96.0155 | 96.0596 |
| Median filter | 15.9871 | 98.1776 | 0.9090 | 98.7174 | 96.4511 | 92.8514 |
| Wiener filter | 0.0896 | 2.0616 | 0.9999 | 99.9998 | 99.9178 | 60.4854 |
| Chebyshev filter | 19.5434 | 99.8199 | 0.9113 | 98.5363 | 76.5784 | 93.2656 |
| Butterworth filter | 19.5087 | 99.7969 | 0.9115 | 98.5428 | 96.1673 | 96.6554 |
| Elliptic filter | 19.4585 | 99.8301 | 0.9114 | 98.5356 | 76.5227 | 93.1759 |
| Savitzky–Golay | 20.3431 | 99.7280 | 0.9115 | 98.6013 | 95.8684 | 97.4314 |
| GRA–ICEEMDAN | 20.8926 | 99.8741 | 0.9115 | 98.5325 | 95.9888 | 98.0180 |

Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SNR: Signal-to-noise ratio.

Models with limited denoising efficacy (e.g., wavelet and Wiener) yielded results nearly identical to the original noisy waveform (Figure 6A), indicating complete failure in noise suppression. Simultaneously, the spectral curves of models, such as elliptic and Butterworth exhibited a noticeable shift in the 0.1–0.7 Hz low-frequency range compared to other models (Figure 6D). This indicates their pronounced low-frequency noise suppression capability, which may play a crucial role in monitoring geomovements and solid tides.

Figure 7 compares the short-time Fourier transform (STFT) spectra of denoised seismic signals recorded by node seismometers under different denoising models. Before the experiment, STFT parameters were uniformly set to ensure comparability among models: Hamming window (256 points), 97.7% overlap, and 512-point Fourier transform. The STFT time–frequency spectrum of the original noisy signal exhibited typical noise contamination characteristics. Across the time axis (0–10 s) and frequency axis (0–100 Hz), the time–frequency energy distribution was diffuse and lacked a distinct structure (Figure 7A), with yellow and orange regions heavily contaminated by noise. Weak blue-highlighted impact structures were visible in the frequency range of 100–250 Hz. Similar spectral patterns were observed in models, such as wavelet, Kalman, moving average, and Wiener (Figure 7B, C, D, and F). The temporal characteristics of the seismic signal were clearly discernible in the STFT spectra of models such as Chebyshev, Butterworth, elliptic, Savitzky–Golay, and GRA–ICEEMDAN (Figure 7G-K), revealing multiple prominent peak contours of the seismic signal. The GRA–ICEEMDAN model uniformly exhibited numerous continuous bright bands around 500 Hz, with impact peak positions showing high amplitude matching to the original signal. In addition, models, such as the median filter exhibited varying degrees of excessive denoising, where the suppression process destroyed original signal details and generated severe artifacts (Figures 7B-E).

Figure 8 presents bar chart comparisons of denoising metrics for ground motion signals recorded by node seismometers under different denoising models. The bars visually illustrate the performance distribution of each model across the overall score and sub-metrics. Notably, the GRA–ICEEMDAN model exhibited significantly higher overall score bars than other models, highlighting its leading position in comprehensive noise suppression performance. This model also demonstrated superior column heights for the SNR improvement and noise suppression rate sub-metrics, confirming its outstanding denoising capability. However, its columns for frequency spectrum similarity and energy retention rate were relatively low, consistent with the model's adjustments to waveforms for removing high-frequency and environmental noise. Nevertheless, the bar chart indicates these metrics remained at high levels (e.g., frequency similarity > 98%, energy retention > 95%).

In contrast, the wavelet and Wiener models' composite score bars were positioned at the lowest level. Yet, their bars for signal correlation, frequency similarity, and energy retention approached or reached the top, starkly contrasting their extremely low bars for SNR improvement and noise suppression rate. This visually reveals these models' severe deficiencies in critical noise suppression metrics, leading to their disadvantageous overall ranking. The multi-subgraph visualization of the bar chart provides an intuitive ranking of each sub-metric in the denoising evaluation. It also clearly reveals the

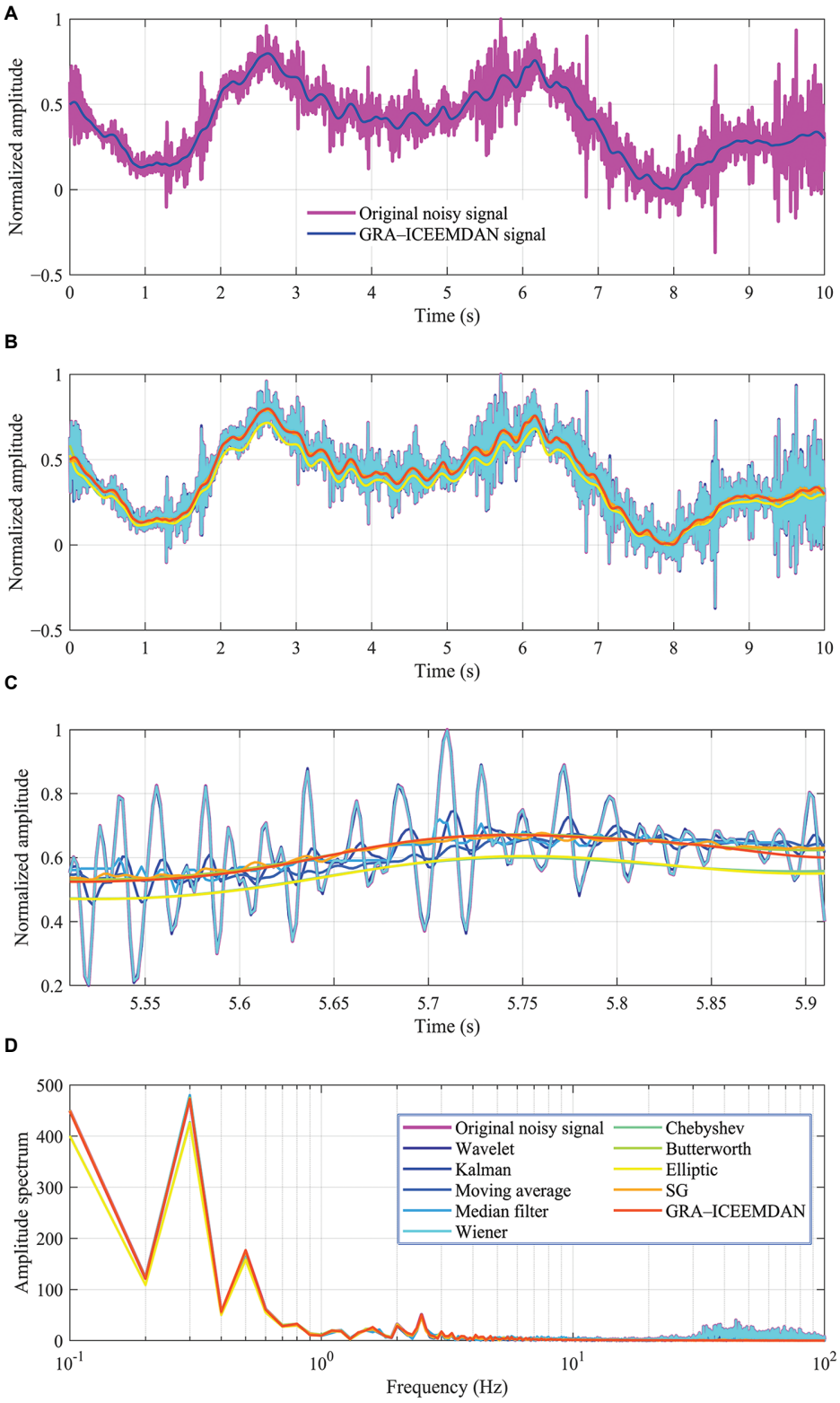



**Figure 6.** Comparison of waveforms for seismic signal denoising results from node seismometers under different denoising models. Comparison of (A) denoising results between the original noisy signal and GRA–ICEEMDAN, (B) results from multiple noise reduction methods, (C) localized amplified waveforms from multiple noise reduction models, and (D) results from multiple denoising models (0.1–100 Hz).

Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SG: Savitzky–Golay.

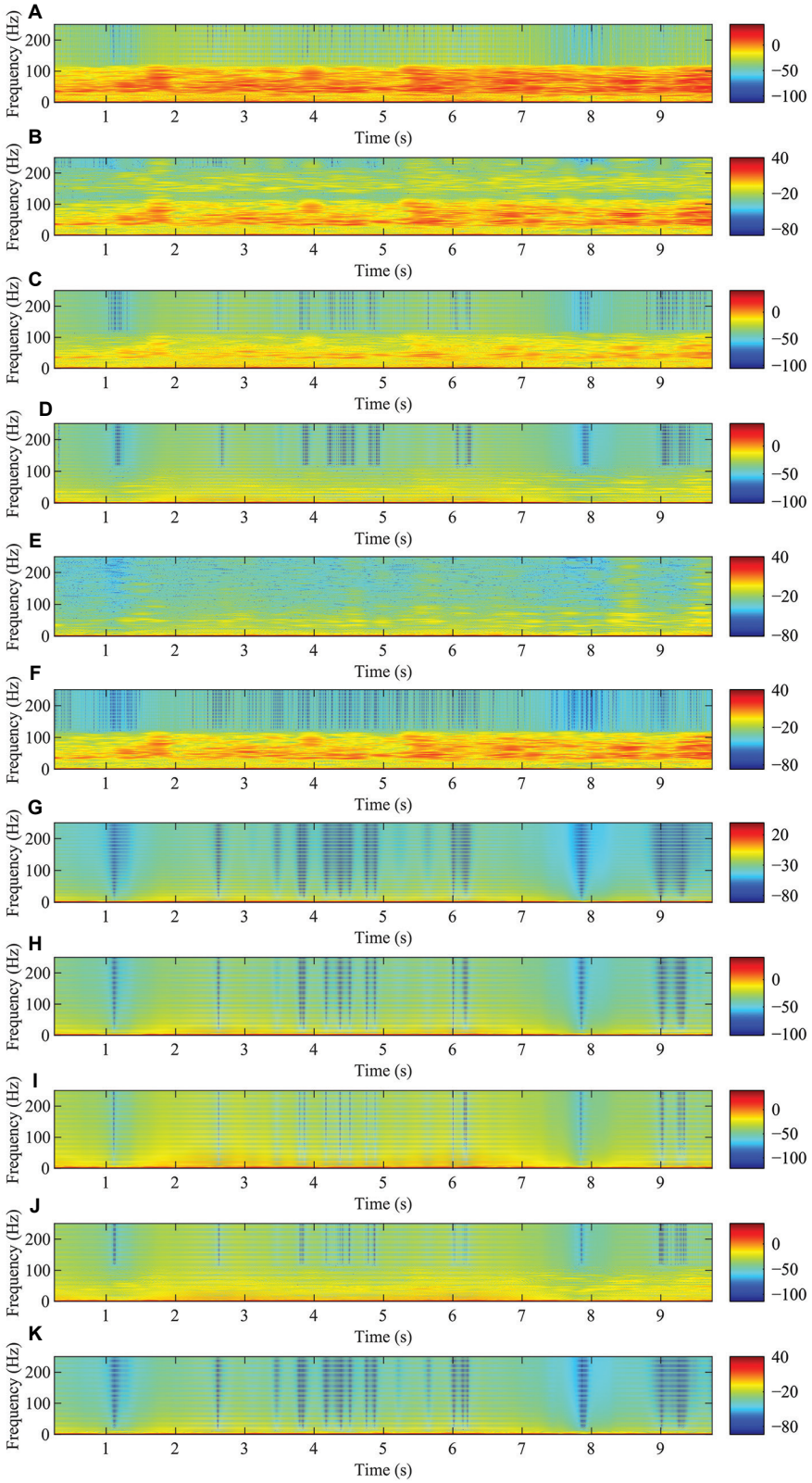



**Figure 7.** Comparison of short-time Fourier transform spectrograms for seismic signal denoising results from node seismometers under different denoising models. (A) Original noisy sound. (B) Wavelet. (C) Kalman. (D) Moving average. (E) Median filter. (F) Wiener. (G) Chebyshev. (H) Butterworth. (I) Elliptic. (J) Savitzky–Golay. (K) Grey relational analysis–improved complete ensemble empirical mode decomposition with adaptive noise signal denoising.

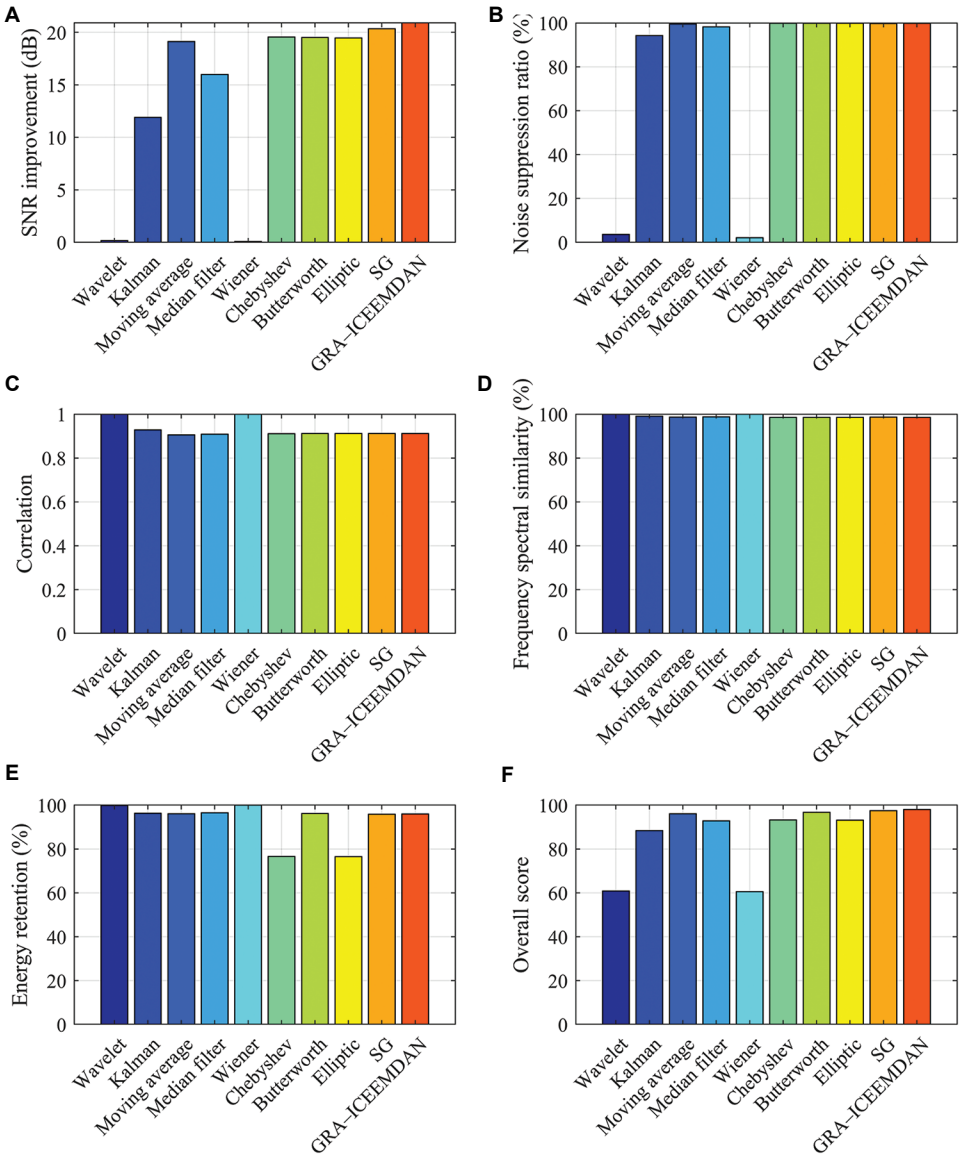



**Figure 8.** Comparison of short-time Fourier transform spectrograms for seismic signal denoising results from node seismometers under different denoising models. (A) Signal-to-noise ratio (SNR) improvement. (B) Noise suppression ratio. (C) Correlation. (D) Frequency spectral similarity. (E) Energy retention. (F) Overall score.
Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SG: Savitzky–Golay.

distribution characteristics of performance differences among models and their strengths and weaknesses across different dimensions, enhancing the objectivity and persuasiveness of the discussion on denoising model results. These detailed metric data are presented in Table 8.

Based on the ranking results by overall score (GRA–ICEEMDAN > Savitzky–Golay > Butterworth filter > Moving average > Chebyshev filter > Elliptic filter > Median filter > Kalman filter > Wavelet > Wiener), the GRA–ICEEMDAN model demonstrated superior overall noise suppression performance with an overall score of 98.0180 (out of 100). Its sub-indicator scores, SNR improvement, noise suppression rate, signal correlation, frequency spectrum similarity, and energy retention rate were 20.892 6 dB, 99.8741%, 0.9115, 98.5325%, and 95.9888%, respectively. All indicators were weighted equally, with a computation time of 3.6484 s. The model also demonstrated balanced performance across individual evaluation metrics, ranking first, first, fourth,

tenth, and seventh in SNR, noise suppression rate, signal correlation, frequency spectrum similarity, and energy retention, respectively, among the 10 denoising models. Most individual metrics were placed within the middle or higher tiers of the ranking cohort. The lower rankings in frequency similarity (98.5325%) and energy retention (95.9888%) stemmed from the GRA–ICEEMDAN model's extensive adjustments to the waveform shape. This approach effectively removed high-frequency components and most environmental noise from the original noisy signal, while still achieving over 98% restoration in the frequency domain.

Some models exhibited poor overall denoising performance, such as the wavelet and Wiener models (overall score < 65), showing significant differences compared to other models. Although wavelet demonstrated outstanding results in certain evaluation metrics (correlation=0.9991, frequency similarity = 99.9874%, energy retention = 99.8499%), with extremely high similarity between pre- and post-denoising results in both time-domain and frequency-domain signals, it failed to accurately assess noise suppression levels (SNR improvement = 0.1605 dB, noise suppression ratio=3.6695%), rendering it uncompetitive in overall rankings. This highlights algorithmic limitations in practical applications, with the Wiener model exhibiting similar deficiencies in noise reduction effectiveness.

However, conducting GRA–ICEEMDAN denoising experiments on natural earthquake events cannot meet the goal of verifying the universality of the new model's denoising. Further denoising experiments on other types of signals recorded by nodal seismographs are needed, which will be analyzed in Section 4.2.

### 4.2. Denoising experiment of ground pulsation signals recorded by node seismometers

The GRA–ICEEMDAN denoising analysis was performed on the unidirectional signal recorded by the node seismometer for the ground pulsation signal. The signal sampling rate was 1,000 Hz. The signal decomposition results, GRA calculation process data, and denoising statistical results are shown in Tables 9-12 and Figures 9-12.

Figure 9 illustrates the nine IMF components obtained through ICEEMDAN decomposition of the ground pulsation signal recorded by the node seismometer. Table 9 lists the characteristic parameter values for each IMF component after ICEEMDAN decomposition. $IMF_1$ to $IMF_4$, obtained from decomposing the noisy signal, all exhibited significant oscillations at relatively high frequencies with small amplitudes. Correspondingly, their sample entropy values all exceeded 0.50, their COEFs were all less than 0.15, and $R^2$ values clustered around 0.01. These waveform characteristics differ markedly from other IMFs, aligning more closely with the random oscillatory nature of environmental noise. Consequently, these components can be identified as system noise requiring removal. The waveform complexity and zero-crossing frequency of $IMF_5$ and $IMF_6$ fell between those of preceding and succeeding IMFs. Their metric values also spanned the range between maximum and minimum values. Whether to include them in the final signal reconstruction requires subsequent ranking and differentiation using the correlation scores from the GRA model. $IMF_7$–$IMF_9$ exhibited smoother waveforms with fewer spikes and pronounced low-frequency characteristics (sample entropy < 0.006, MSE > 0.2), closely resembling standard sine wave curves. These are presumed to be effective components of the original signal. Notably, $IMF_7$ exhibited high consistency with the

**Table 9. Characteristic parameter values of each IMF after ICEEMDAN decomposition of the ground pulsation signals recorded by nodal seismographs**

| IMF | Correlation coefficient | Sample Entropy | CS | $R^2$ | JSD | RMSE | MAE | MAPE | Adj-$R^2$ | MI |
|---|---|---|---|---|---|---|---|---|---|---|
| $IMF_1$ | 0.1028 | 1.9775 | 0.1029 | 0.0101 | 0.3769 | 77.9736 | 68.8421 | 1.0463e+16 | 0.0104 | 0.6487 |
| $IMF_2$ | 0.1321 | 1.3067 | 0.1321 | 0.0161 | 0.3682 | 77.7401 | 68.7084 | 1.2036e+16 | 0.0173 | 0.6759 |
| $IMF_3$ | 0.1205 | 0.6596 | 0.1206 | 0.0137 | 0.3659 | 77.8339 | 68.7215 | 1.3078e+16 | 0.0144 | 0.6927 |
| $IMF_4$ | 0.1026 | 0.5525 | 0.1026 | 0.0100 | 0.3630 | 77.9782 | 68.7346 | 1.1521e+16 | 0.0104 | 0.6918 |
| $IMF_5$ | 0.1502 | 0.2163 | 0.1503 | 0.0214 | 0.3657 | 77.5308 | 68.5060 | 1.6317e+16 | 0.0224 | 0.6999 |
| $IMF_6$ | 0.1681 | 0.0798 | 0.1679 | 0.0260 | 0.3238 | 77.3465 | 68.5524 | 1.3125e+16 | 0.0281 | 0.7072 |
| $IMF_7$ | 0.9209 | 0.0059 | 0.9154 | 0.8213 | 0.0414 | 33.1294 | 26.5277 | 3.1722e+16 | 0.8480 | 0.7182 |
| $IMF_8$ | 0.7955 | 0.0046 | 0.7810 | 0.2680 | 0.1198 | 67.0541 | 57.5983 | 1.6269e+16 | 0.6328 | 0.7166 |
| $IMF_9$ | 0.2771 | 0.0027 | 0.2272 | 0.0500 | 0.4571 | 76.3866 | 66.3474 | 2.1118e+16 | 0.0767 | 0.7085 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

original signal (COEF = 0.9209, $R^2$ = 0.8213, JSD = 0.0414, Adj-$R^2$ = 0.8480, MAE = 26.5277), and can be regarded as the primary noise-free component extracted from the original noisy signal.

After obtaining the ICEEMDAN decomposition results of the ground pulsation signals recorded by the nodal seismographs (Figure 9), the 10 characteristic index values of each of the above IMFs were calculated, as shown in Table 9.

The IMF quality evaluation indicator matrix was derived from Table 9. After undergoing necessary processing steps—including normalization, two-stage least squares/maximum differences calculation, grey relational coefficient computation, subjective assignment of indicator weights, and IMF correlation analysis—the final quality ranking results for IMF components were obtained. The specific calculation results are shown in Tables 10 and 11. Table 10 presents the comparative sequences and reference sequence values for each IMF of the ground pulsation signals recorded by the node seismometer, while Table 11 displays the GRA COEFs and correlation values for each IMF of the ground pulsation signals recorded by the node seismometer.

**Table 10. Comparison and reference series values of ground vibration signals recorded by nodal seismographs for each IMF**

| Index | The IMF evaluated | | | | | | | | | Ideal IMF |
|---|---|---|---|---|---|---|---|---|---|---|
| | $IMF_1$ | $IMF_2$ | $IMF_3$ | $IMF_4$ | $IMF_5$ | $IMF_6$ | $IMF_7$ | $IMF_8$ | $IMF_9$ | |
| Correlation coefficient | 0.0002 | 0.0360 | 0.0218 | 0 | 0.0581 | 0.0801 | 1 | 0.8467 | 0.2133 | 1 |
| Sample entropy | 0 | 0.3396 | 0.6673 | 0.7215 | 0.8918 | 0.9609 | 0.9983 | 0.9990 | 1 | 1 |
| CS | 0.0004 | 0.0363 | 0.0221 | 0 | 0.0586 | 0.0803 | 1 | 0.8346 | 0.1532 | 1 |
| $R^2$ | 0.0001 | 0.0074 | 0.0045 | 0 | 0.0139 | 0.0196 | 1 | 0.3179 | 0.0493 | 1 |
| JSD | 0.1927 | 0.2136 | 0.2191 | 0.2260 | 0.2195 | 0.3205 | 1 | 0.8113 | 0 | 1 |
| RMSE | 0.0001 | 0.0053 | 0.0032 | 0 | 0.0099 | 0.0140 | 1 | 0.2435 | 0.0354 | 1 |
| MAE | 0 | 0.0031 | 0.0028 | 0.0025 | 0.0079 | 0.0068 | 1 | 0.2657 | 0.0589 | 1 |
| MAPE | 1 | 0.9260 | 0.8770 | 0.9502 | 0.7246 | 0.8747 | 0 | 0.7269 | 0.4987 | 1 |
| Adj-$R^2$ | 0.0005 | 0.0082 | 0.0047 | 0 | 0.0143 | 0.0212 | 1 | 0.7429 | 0.0791 | 1 |
| MI | 0 | 0.3917 | 0.6337 | 0.6207 | 0.7369 | 0.8419 | 1 | 0.9780 | 0.8613 | 1 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

**Table 11. GRA correlation coefficients and correlation values of each IMF in the ground pulsation signals recorded by nodal seismographs**

| Evaluation indicators | The IMFs evaluated | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $IMF_1$ | $IMF_2$ | $IMF_3$ | $IMF_4$ | $IMF_5$ | $IMF_6$ | $IMF_7$ | $IMF_8$ | $IMF_9$ |
| Correlation coefficient | 0.3333 | 0.3415 | 0.3382 | 0.3333 | 0.3467 | 0.3521 | 1 | 0.7653 | 0.3885 |
| Sample entropy | 0.3333 | 0.4309 | 0.6004 | 0.6423 | 0.8221 | 0.9275 | 0.9966 | 0.9980 | 1 |
| CS | 0.3334 | 0.3416 | 0.3383 | 0.3333 | 0.3468 | 0.3521 | 1 | 0.7515 | 0.3712 |
| $R^2$ | 0.3333 | 0.3349 | 0.3343 | 0.3333 | 0.3364 | 0.3377 | 1 | 0.4229 | 0.3446 |
| JSD | 0.3824 | 0.3887 | 0.3903 | 0.3924 | 0.3904 | 0.4239 | 1 | 0.7260 | 0.3333 |
| RMSE | 0.3333 | 0.3345 | 0.3340 | 0.3333 | 0.3355 | 0.3364 | 1 | 0.3979 | 0.3414 |
| MAE | 0.3333 | 0.3340 | 0.3339 | 0.3338 | 0.3351 | 0.3348 | 1 | 0.4050 | 0.3469 |
| MAPE | 1 | 0.8711 | 0.8025 | 0.9094 | 0.6448 | 0.7997 | 0.3333 | 0.6467 | 0.4993 |
| Adj-$R^2$ | 0.3333 | 0.3351 | 0.3343 | 0.3333 | 0.3365 | 0.3381 | 1 | 0.6604 | 0.3519 |
| Mutual Information | 0.3333 | 0.4511 | 0.5772 | 0.5686 | 0.6552 | 0.7597 | 1 | 0.9579 | 0.7829 |
| Correlation | 0.4049 | 0.4164 | 0.4384 | 0.4513 | 0.4550 | 0.4963 | 0.9330 | 0.6732 | 0.4760 |

Abbreviations: Adj-$R^2$: Adjusted coefficient of determination; CS: Cosine similarity; IMF: Intrinsic mode function; JSD: Jensen–Shannon divergence; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MI: Mutual information; $R^2$: Coefficient of determination; RMSE: Root mean square error.

**Table 12. Comparison of nodal seismograph signal denoising effects under different models**

| Denoising model | SNR improvement (db) | Noise suppression ratio (%) | Correlation | Frequency similarity (%) | Energy retention (%) | Overall score |
|---|---|---|---|---|---|---|
| Wavelet | 1.3157 | 25.4934 | 0.9977 | 99.9124 | 98.9379 | 66.1399 |
| Kalman filter | 10.8861 | 90.7295 | 0.9790 | 99.2136 | 95.8323 | 87.6389 |
| Moving average | 14.7322 | 95.4371 | 0.9552 | 98.4321 | 94.4711 | 91.5354 |
| Median filter | 14.1609 | 95.2443 | 0.9770 | 98.9580 | 95.0623 | 91.5796 |
| Wiener filter | 0.0759 | 1.6942 | 0.9999 | 99.9999 | 99.9272 | 60.4002 |
| Chebyshev filter | 16.1413 | 96.8667 | 0.9697 | 98.5302 | 75.7184 | 90.3033 |
| Butterworth filter | 16.2441 | 95.8946 | 0.9713 | 98.6158 | 93.9739 | 93.4027 |
| Elliptic filter | 16.1725 | 96.6147 | 0.9709 | 98.6003 | 75.7323 | 90.3244 |
| Savitzky–Golay | 15.1727 | 95.9182 | 0.9761 | 98.9484 | 94.6695 | 92.6313 |
| GRA–ICEEMDAN | 16.3523 | 96.3398 | 0.9774 | 98.8683 | 94.6693 | 93.9056 |

Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SNR: Signal-to-noise ratio.

The comparison and reference sequences in Table 10 can be used to construct the GRA COEFs between 10 evaluation indicators and each IMF (Table 11). Based on the balanced indicator weight vector {0.1,0.1,0.1,0.1,0.1,0.1 ,0.1,0.1,0.1,0.1,0.1} and the GRA COEFs, the grey weighted correlation degree of each IMF was obtained using the arithmetic average (Table 11).

$IMF_7$ achieved the highest COEF (value of 1) in 8 out of 10 indicators, indicating its most significant correlation with the original signal and the highest overall score (0.9330). $IMF_1$ and $IMF_9$ achieved the best results in MAPE and sample entropy indicators, respectively. The GRA comprehensive evaluation ranking is as follows: $IMF_7$ > $IMF_8$ (0.6732) > $IMF_6$ (0.4963) > $IMF_9$ (0.4760) > $IMF_5$ (0.4550) > $IMF_4$ (0.4513) > $IMF_3$ (0.4384) > $IMF_2$ (0.4164) > $IMF_1$ (0.4049). Although $IMF_5$, $IMF_6$, and $IMF_8$ did not achieve individual optimal values, their multiple COEFs exceeded 0.6, indicating high correlation with the signal and warranting retention. The remaining IMFs were discarded due to generally low correlation levels. The model selected the top five IMF components ($IMF_7$, $IMF_8$, $IMF_6$, $IMF_9$, $IMF_5$) for linear reconstruction, with the denoising results shown in Figure 10.

To evaluate the superiority of the GRA–ICEEMDAN model among comparable signal denoising methods, nine alternative models were selected for comparison: wavelet threshold denoising, median filter, Wiener filter, Savitzky–Golay filter, elliptic filter, Butterworth filter, Kalman filter, Chebyshev filter, and moving average filter. The evaluation framework incorporated multiple quantitative metrics across dimensions, including SNR improvement, noise suppression ratio, correlation, frequency similarity, energy retention, and overall score. The denoising results for each model are shown in Figure 10.

Except for the wavelet and Wiener models, other models produced positive effects in denoising geopotential signals, achieving a certain degree of smoothing and high time-frequency fidelity (Figure 10B and D), though with notable differences in waveform processing. Local magnified views revealed that the GRA–ICEEMDAN model exhibited superior phase consistency and better preservation of peak oscillation trends. Other models (e.g., moving average) directly filtered out local detailed features near peaks. The smoother curves obtained come at the cost of sacrificing more local effective components and phase information, which hinders the capture of signal details during subsequent data processing stages.

Models with limited denoising effectiveness (e.g., wavelet and Wiener) largely overlapped with the original noisy waveform (Figure 10A), indicating suboptimal noise suppression. In addition, the spectral curves of models, such as elliptic and Butterworth exhibited a certain degree of shift in the 0.1–0.4 Hz low-frequency range compared to other models (Figure 10D). This partially confirms that some models sacrifice local temporal features during denoising, thereby limiting their applicability in high-precision seismic monitoring and geophysical exploration.

Figure 11 compares the STFT spectra of denoised ground motion signals recorded by node seismometers under different denoising models. Before the experiment, STFT parameters were uniformly set to ensure comparability among models: Hamming window (256 points), 97.7% overlap, and 512-point Fourier transform. The STFT time–frequency spectrum of the original noisy signal exhibited

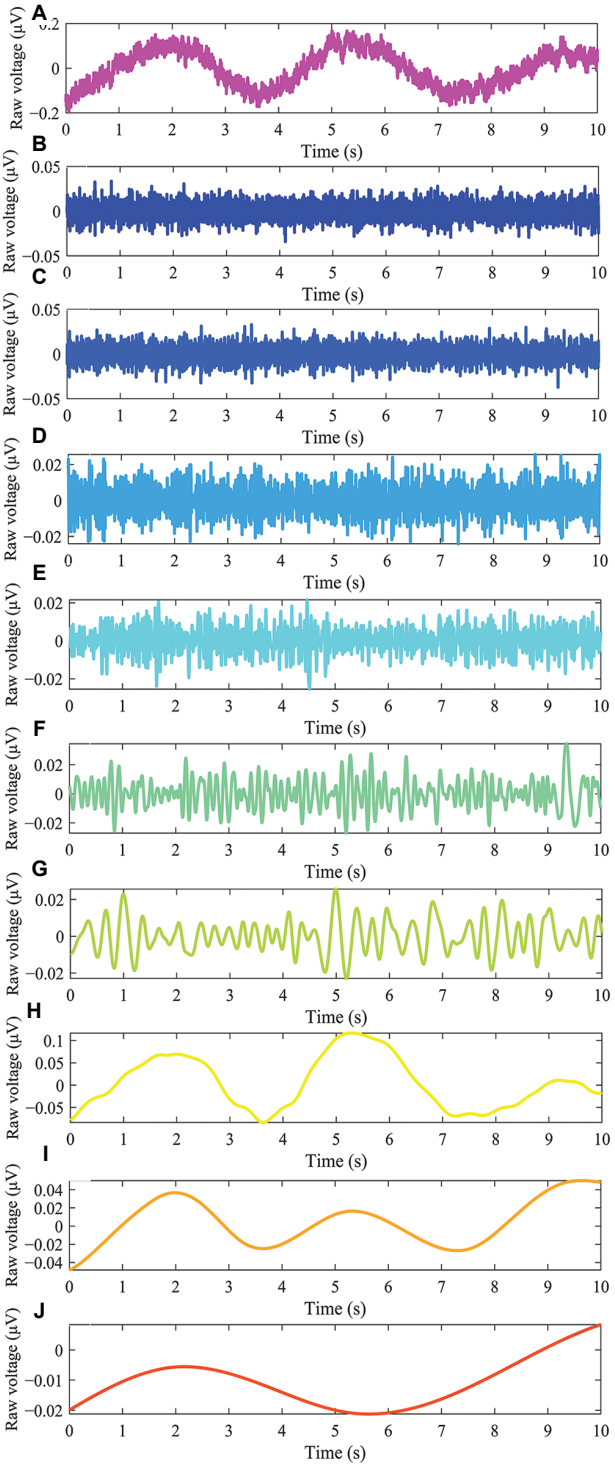



**Figure 9.** Results of the ground pulsation signals recorded by the nodal seismographs after ICEEMDAN decomposition. (A) Original noisy signal. (B) $IMF_1$. (C) $IMF_2$. (D) $IMF_3$. (E) $IMF_4$. (F) $IMF_5$. (G) $IMF_6$. (H) $IMF_7$. (I) $IMF_8$. (J) $IMF_9$.
Abbreviation: ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; IMF: Intrinsic mode function.

typical noise contamination characteristics. Across the entire time axis (0–10 s) and frequency axis (0–400 Hz), the time–frequency energy distribution was diffuse and lacked a distinct structure (Figure 11A). In the region above the background noise (yellow and orange) at 400–500 Hz, the signal was severely contaminated by noise. Similar spectral patterns were observed for the wavelet model and Wiener model (Figure 11B and F). The temporal characteristics of the geopulse signal WERE clearly visible in the STFT spectra of models such as Chebyshev, Butterworth, elliptic, Savitzky–Golay, and GRA–ICEEMDAN (Figure 11G-K). Multiple prominent peak contours of the geodetic signal were discernible, while the GRA–ICEEMDAN model uniformly displayed numerous continuous bright bands at 500 Hz. The positions of impact peaks exhibited high amplitude, matching the original signal. Furthermore, models such as the median filter, the moving average, and the Kalman filter exhibited varying degrees of excessive noise reduction. The suppression process destroyed original signal details and produced severe artifacts (Figures 11B-E).

Figure 12 presents bar chart comparisons of denoising metrics for ground motion signals recorded by node seismometers under different denoising models. The bars visually illustrate the performance comparison across comprehensive scores and sub-metrics for each model. Notably, the GRA–ICEEMDAN model exhibited significantly higher comprehensive scores than other models, confirming its superior time–frequency balance across multiple evaluation metrics and its leading overall noise reduction performance. In contrast, the wavelet and Wiener models exhibited notably lower composite score bars, creating a stark contrast with their relatively high bars in certain metrics (e.g., signal correlation, frequency similarity). This further highlights their shortcomings in key noise suppression metrics (SNR improvement, noise suppression ratio), resulting in their disadvantageous position in the overall denoising evaluation. The multi-subgraph visualization of the bar chart provided an intuitive ranking of performance across sub-metrics in the denoising evaluation. It also clearly revealed the distribution characteristics of performance differences between models and their strengths and weaknesses across various dimensions, enhancing the objectivity and persuasiveness of the discussion on denoising model results. These detailed metric data are presented in Table 12.

Based on the overall score ranking (GRA–ICEEMDAN > Butterworth filter > Savitzky–Golay > Median filter > Moving average > Elliptic filter > Chebyshev filter > Kalman filter > Wavelet > Wiener), the GRA–

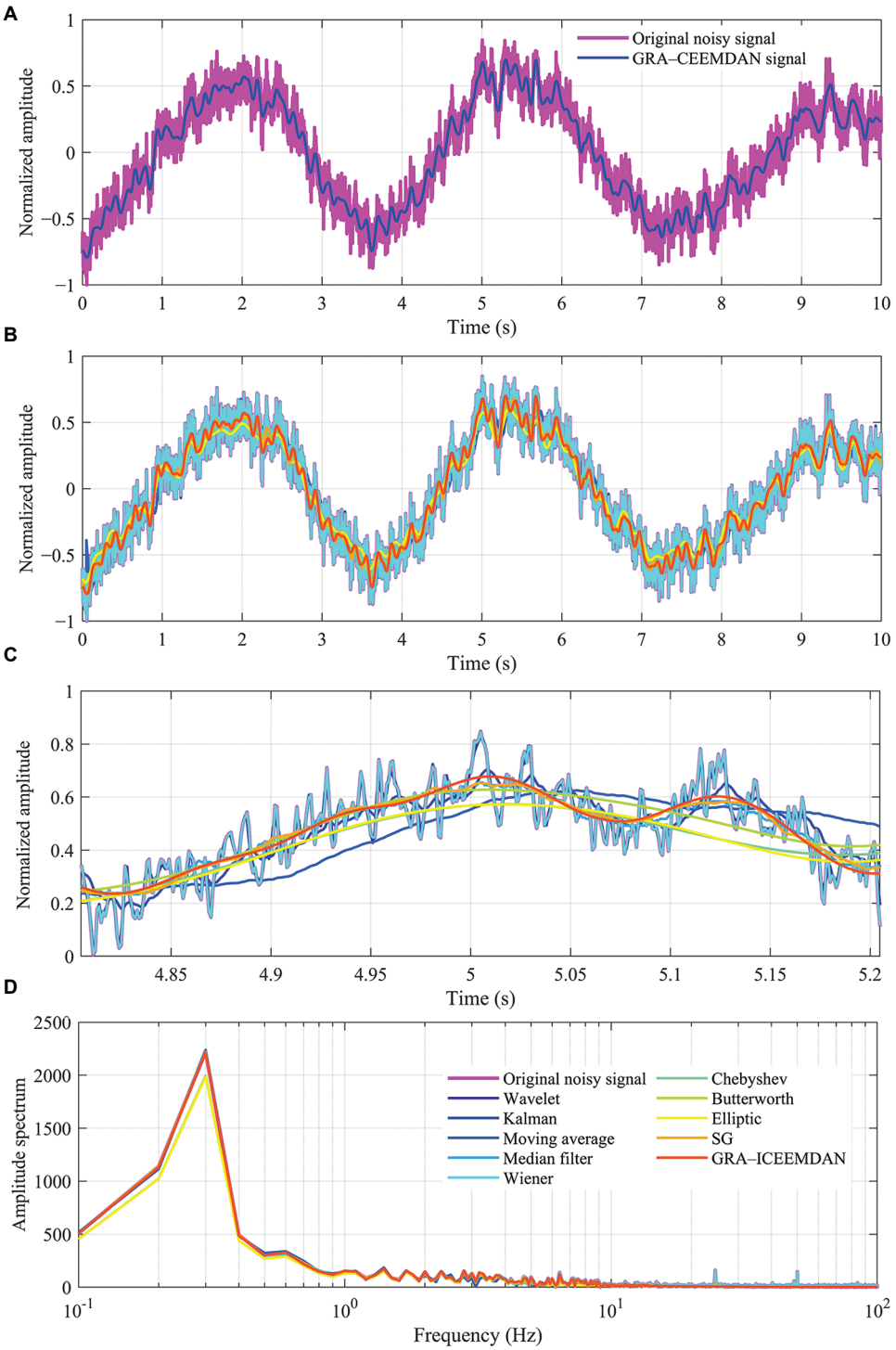



**Figure 10.** GRA–ICEEMDAN denoising results of ground vibration signals recorded by nodal seismographs. Comparison of (A) denoising results between the original noisy signal and GRA–ICEEMDAN, (B) results from multiple noise reduction methods, (C) localized amplified waveforms from multiple noise reduction models, and (D) results from multiple denoising models (0.1–100 Hz).

Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SG: Savitzky–Golay.

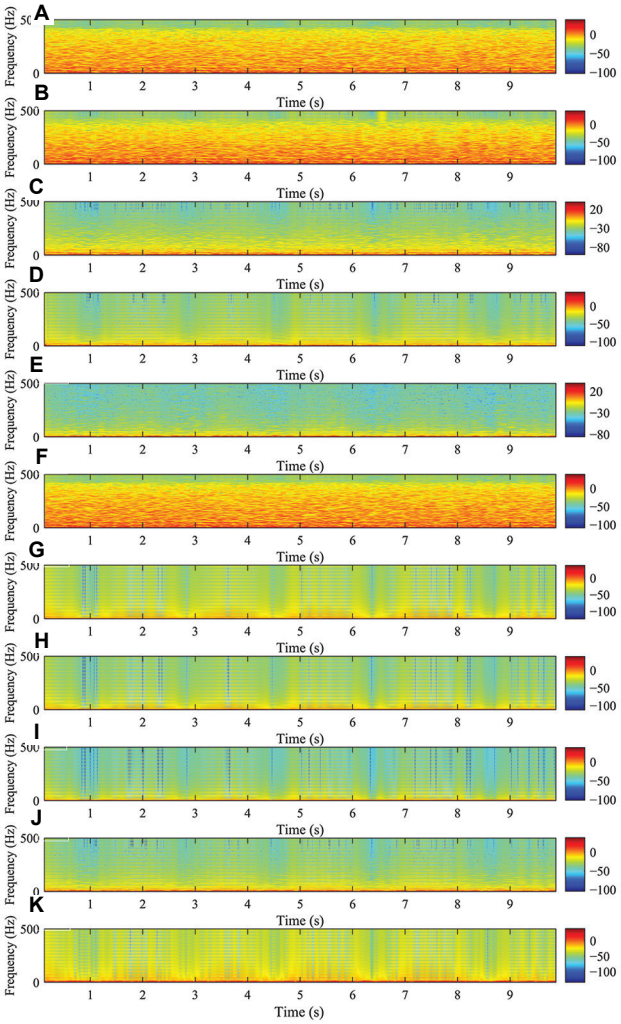



**Figure 11.** Comparison of short-time Fourier transform spectrograms for ground pulse signal denoising results from node seismometers under different denoising models. (A) Original noisy sound. (B) Wavelet. (C) Kalman. (D) Moving average. (E) Median filter. (F) Wiener. (G) Chebyshev. (H) Butterworth. (I) Elliptic. (J) Savitzky–Golay. (K) Grey relational analysis–improved complete ensemble empirical mode decomposition with adaptive noise signal denoising.

ICEEMDAN model demonstrated superior overall noise suppression performance. Its comprehensive score is 93.9056 (out of 100), with sub-metric scores as follows: SNR improvement: 16.3523 dB, noise suppression rate: 96.3398%, signal correlation: 0.9774, frequency spectrum similarity: 98.8683%, and energy retention rate: 94.6693%. The computation time was 12.9703 s. The model also demonstrated balanced performance across individual evaluation metrics, ranking first, third, fourth, sixth, and sixth in SNR, noise suppression rate, signal correlation, frequency spectrum similarity, and energy retention, respectively, among the 10 denoising models. All individual metrics were at or above the mid-range level in the ranking cohort, with no instances of poor "specialization" or concentrated performance.

Some models exhibited poor overall noise reduction performance, such as the wavelet and Wiener models (overall score <70), showing significant differences compared to other models. Although the wavelet model demonstrated outstanding results in certain evaluation metrics (correlation = 0.9977, frequency similarity = 99.9124%, energy retention = 98.9379%), with extremely high similarity between pre- and post-denoising results in both time-domain and frequency-domain signals, it failed to accurately assess noise suppression levels (SNR improvement = 1.3157 dB, noise suppression ratio = 25.4934%), rendering them uncompetitive in overall rankings. This highlights algorithmic shortcomings in practical applications, with the Wiener model exhibiting similar deficiencies.

## 5. Discussion

In addition to the SNR, the practical application effectiveness of denoising models requires consideration of more complex aspects, such as model robustness and algorithm robustness. To further evaluate the multi-round repeatable denoising robustness of the GRA–ICEEMDAN model, a cyclic denoising experiment was designed for nodal seismograph-recorded ground pulsation data with 20 rounds. The evaluation metrics for the denoising model included COEF, absolute difference of distribution entropy, reconstructed SNR, and MSE. Statistical indicators such as mean, standard deviation, range, and median were used to measure the consistency of the multi-round denoising results. The experimental results are shown in Figure 13 and Table 13.

Figure 13 shows the denoising results of 20 rounds of GRA-ICEEMDAN recorded ground pulsation signals by nodal seismographs, and Table 13 shows the statistical table of the denoising results of 20 rounds of nodal seismograph recorded ground pulsation signals. The denoising results of each time were highly correlated with the original waveform (mean >0.97, standard deviation <0.01), and no special cases of distortion or deformation occurred. The reconstructed SNR was greater than 10 dB, indicating that the GRA–ICEEMDAN model has a good random suppression effect on a large amount of environmental noise. The absolute value of the distribution entropy difference was all below 0.05, reflecting that the overall consistency of the waveform period before and after denoising by the GRA–ICEEMDAN model was high.

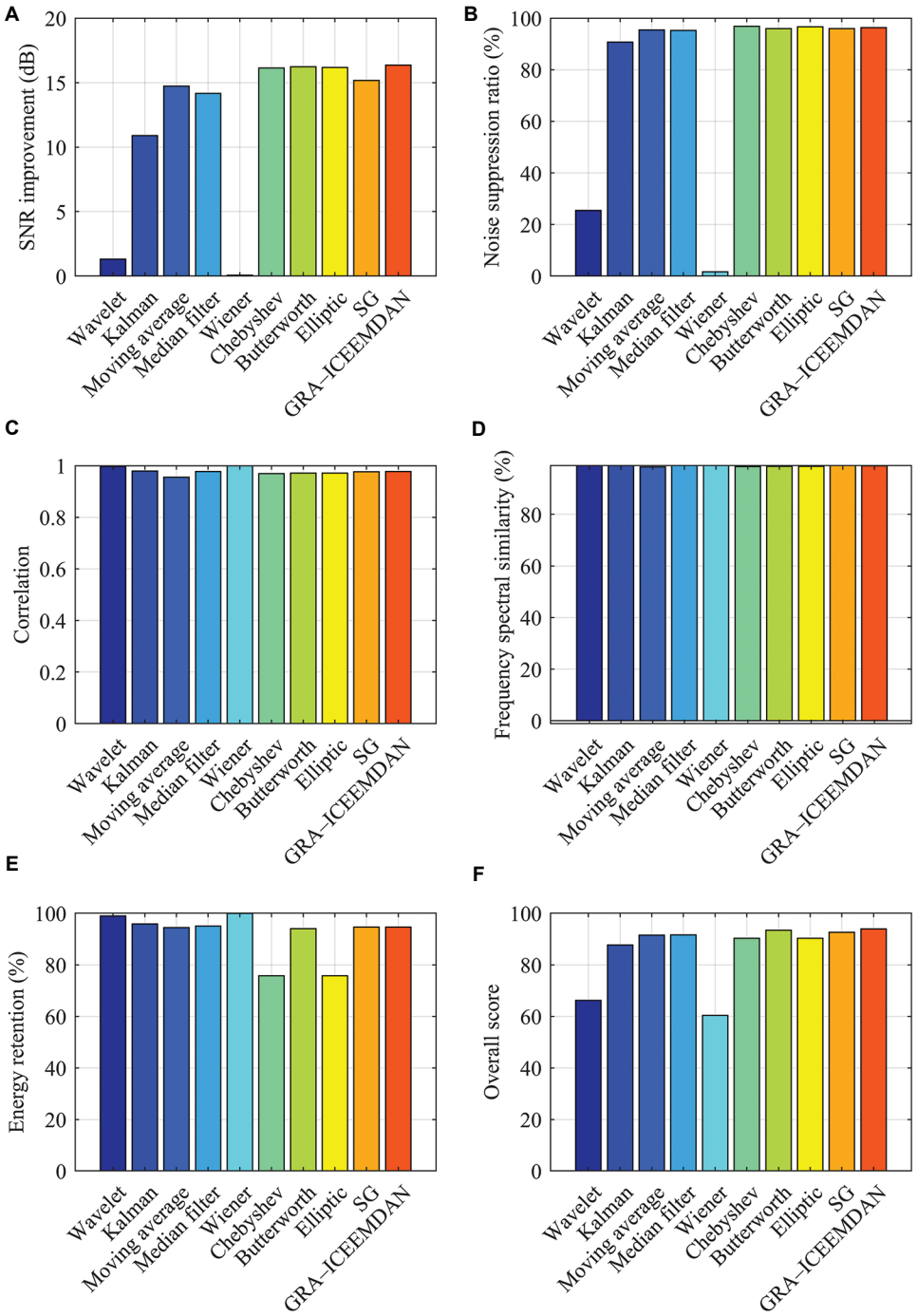



**Figure 12.** Bar chart comparison of denoising metrics for ground pulse signal recorded by node seismographs under different denoising models. (A) Signal-to-noise ratio (SNR) improvement. (B) Noise suppression ratio. (C) Correlation. (D) Frequency spectral similarity. (E) Energy retention. (F) Overall score.
Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising; SG: Savitzky–Golay.

The repeatability test results of the GRA–ICEEMDAN model in this section verified the stability of its denoising effect, improved the simulation test of the model, and largely demonstrated that the noise processing process of GRA combined with ICEEMDAN is feasible and reliable.

**Table 13. Statistical table of 20 rounds of noise reduction results for ground pulsation signals recorded by nodal seismographs**

| Evaluation metrics | Mean | STD | Range | Median |
|---|---|---|---|---|
| Correlation coefficient | 0.9720 | 0.0083 | 0.0323 | 0.9770 |
| Mean square error | 366.9937 | 115.6653 | 374.5581 | 288.7540 |
| Absolute difference in distribution entropy | 0.0336 | 0.0074 | 0.0308 | 0.0348 |
| Reconstructed signal-to-noise ratio (dB) | 12.4271 | 1.2874 | 3.8006 | 13.2809 |

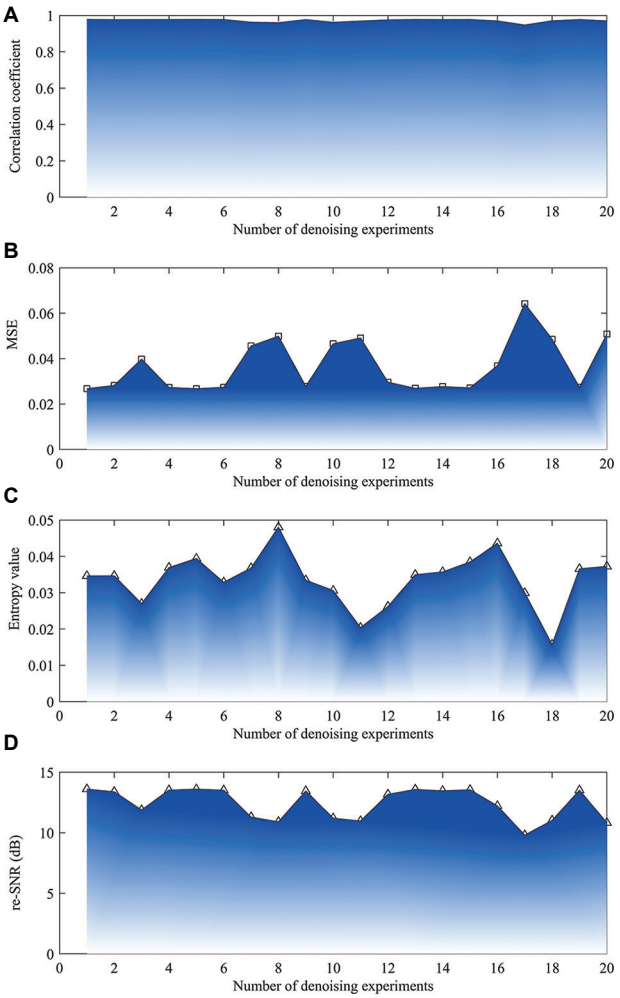



**Figure 13.** Twenty rounds of GRA–ICEEMDAN denoising results of ground pulsation signals recorded by nodal seismographs. (A) Correlation coefficient. (B) Mean square error. (C) Absolute value of the distribution entropy difference. (D) Reconstructed signal-to-noise ratio.
Abbreviations: GRA: Grey relational analysis; ICEEMDAN: Improved complete ensemble empirical mode decomposition with adaptive noise signal denoising.

## 6. Conclusion

This study proposed a joint denoising model based on ICEEMDAN and GRA to suppress random noise in various observation signals from nodal seismographs. Across systematic experiments on simulated signals, actual ground pulsation signals, and seismic event signals, as well as comparisons with traditional methods, such as Kalman filtering, Chebyshev filtering, Savitsky–Golay filtering, and Butterworth filtering, the following conclusions are drawn:

(i) Denoising effectiveness: This model achieved an excellent balance between preserving effective signals and suppressing random noise. In simulation experiments, it showed the greatest improvement in reconstructed SNR and the lowest MSE. In ground pulsation signal processing, it effectively smoothed the noise background and highlighted potential weak reflection information. In seismic event applications, it achieved the highest waveform fidelity for key phases, such as P-waves, effectively avoiding amplitude loss or phase distortion caused by excessive denoising.

(ii) Method advantages: ICEEMDAN decomposition effectively overcame the mode aliasing problem of traditional EMD, laying a solid foundation for an accurate selection of subsequent components. The core innovation lies in introducing GRA as the discrimination criterion for IMF components. By calculating the geometric similarity between each IMF and the noisy original signal, it can adaptively and precisely distinguish between components dominated by noise and those dominated by effective signals, overcoming the rigid judgment defects of traditional methods based on COEFs or fixed thresholds.

(iii) Robustness and universality: Three sets of experiments of different types consistently showed that this method has good adaptability and robustness for seismic signals with different SNRs and characteristics, verifying its great potential as a high-fidelity, adaptive denoising tool in actual seismic data processing.

In summary, the ICEEMDAN–GRA integration technique provides a balanced and effective solution for high-quality pre-processing of node seismometer data, particularly suitable for seismic monitoring and geophysical exploration projects involving small to medium-sized samples. Although this study has achieved some anticipated results, the following aspects warrant further exploration and refinement:

(i) Parameter adaptive optimization: While parameters, such as ICEEMDAN decomposition noise intensity,

ensemble size, and maximum iteration count are effective in the current model, they rely to some extent on empirical settings. Future work should explore integrating intelligent optimization algorithms (e.g., particle swarm optimization, genetic algorithms) with this research model to achieve fully adaptive determination of key parameters, thereby further enhancing the model's automation level and universality.

(ii) Graphics Processing Unit parallel computing: Each noise addition decomposition in ICEEMDAN is inherently independent, presenting a natural parallelism opportunity. Calculations for different IMFs and metrics within GRA can also be parallelized, making the algorithm highly amenable to parallel acceleration on Graphics Processing Units or multi-core Central Processing Units. Near-linear speedup is anticipated.

(iii) High-performance computing and cloud computing: For ultra-large-scale seismic datasets, data can be partitioned and processed in parallel across clusters, thereby reducing the absolute computational time of the model presented in this study.

## Acknowledgments

None.

## Funding

This research was financially supported by Mahasarakham University; the Scientific Research Fund from the Institute of Seismology, CEA and National Institute of Natural Hazards, Ministry of Emergency Management of China (No. IS202226322, No.IS202436357); the Open Fund of Wuhan Gravitation and Solid Earth Tides, National Observation and Research Station, (No.WHYWZ202406); the Spark Program of Earthquake Technology of CEA (No.XH24025YC, XH25019C); the 2025 Hubei Provincial Key Laboratory Targeted Commissioned Project (No.2025CSA083); research grants from the National Institute of Natural Hazards, Ministry of Emergency Management of China (Grant number: ZDJ2024-31); the 2025 Doctoral Special Support Program Project of Chengdu Jincheng College (NO.2025JCKY(B)0018); and the Mathematics and Finance Research Center Project of Dazhou Social Science Federation Key Research Base (No. SCMF202505).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Cong Pang, Tianwen Zhao, Sirui Liu, Xingxing Li, Ya Xiang

*Formal analysis:* Cong Pang, Guoqing Chen, Tianwen Zhao, Yuxuan Liang, Ya Xiang

*Investigation:* Tianwen Zhao, Sirui Liu, Yuxuan Liang, Xingxing Li, Piyapatr Busababodhin

*Methodology:* Cong Pang, Tianwen Zhao, Guoqing Chen, Piyapatr Busababodhin

*Writing–original draft:* Cong Pang, Tianwen Zhao, Guoqing Chen, Piyapatr Busababodhin

*Writing–review & editing: Cong Pang,* Tianwen Zhao, Guoqing Chen, Piyapatr Busababodhin

## 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 upon reasonable request.

## References

1. Baker B, Holt MM, Pankow KL, Koper KD, Farrell J. Monitoring the 2020 Magna, Utah, earthquake sequence with nodal seismometers and machine learning. *Seismol Res Lett*. 2021;92(2A):787-801.

   doi: 10.1785/0220200316

2. Xiong C, Deng Y, Wang L, Ye X, Zhang Y. High-resolution monitoring of Urban activities with nodal seismometers in the Guangdong-Hong Kong-Macao greater bay area, China. *Earth Space Sci*. 2025;12(4):e2024EA004088.

   doi: 10.1029/2024EA004088

3. Nayak A, Verónica RT, Jonathan AF, *et al*. Nodal and broadband seismometer complement to the imperial valley dark fiber DAS array. *Seismol Res Lett*. 2023;94(6):2852-2867.

   doi: 10.1785/0220230081

4. Gaishan Z, Zhanxiang HE. Comprehensive view of the development of nodal seismic technology. *Geophys Prospect Petrol*. 2024;63(4):718-734.

   doi: 10.12431/issn.1000-1441.2024.63.04.002

5. Li YE, Nilot EA, Zhao Y, Fang G. Quantifying Urban activities using nodal seismometers in a heterogeneous Urban space. *Sensors (Basel)*. 2023;23(3):1322.

   doi: 10.3390/s23031322

6. Abbas A, Cox BR, Tran KT, Corey I, Dawadi N. An open-access data set of active-source and passive-wavefield DAS and nodal seismometer measurements at the newberry florida site. *Seismol Res Lett*. 2024;95(2A):1082-1098.

   doi: 10.1785/0220230216

7. Liu S, Yan A, Huang S. Seismic data denoising based on DC-PCNN image fusion in NSCT domain. *IEEE Geosci Remote Sens Lett*. 2024;21:1-5.

   doi: 10.1109/LGRS.2024.3369727

8. Liang C, Lin H, Ma H. Reinforcement learning-based

denoising model for seismic random noise attenuation. *IEEE Trans Geosci Remote Sens*. 2023;61:1-17.

doi: 10.1109/TGRS.2023.3268718

9. Li F, Liu H, Wang W, Ma J. Swin transformer for seismic denoising. *IEEE Geosci Remote Sens Lett*. 2024;21:1-5.

doi: 10.1109/LGRS.2024.3358234

10. Cao M, Hua Y, Zhang J, Zhang H, Ma F, Ji P. Denoising method for ultrasonic echo signal of mining well logging instrument based on NRBO-ICEEMDAN wavelet thresholding. *IEEE Sens J*. 2025;25(15):29700-29710.

doi: 10.1109/JSEN.2025.3578599

11. Chen X, Chen XY, Zhao LY, *et al*. Feature-adaptive self-supervised leaning for microthrust measurement signals' blind denoising. *IEEE Sens J*. 2025;25(1):1015-1028.

doi: 10.1109/JSEN.2024.3493104

12. Qiao J, Ai W, Hu X, Liu M, Hu S. An identification method of melting layer using the covariance wavelet transform based on GPM-DPR observations. *Earth Space Sci*. 2022;9(3):e2021EA002103.

doi: 10.1029/2021EA002103

13. Sun M, Wu J, Lu Y, Yu F, Zhou H. Engineering safety-oriented blasting-induced seismic wave signal processing: An EMD endpoint suppression method based on multi-scale feature. *Sensors (Basel)*. 2025;25(13):4194.

doi: 10.3390/s25134194

14. Zhao Y, Zhong Z, Li Y, *et al*. Ensemble empirical mode decomposition and stacking model for filtering borehole distributed acoustic sensing records. *Geophysics*. 2023;88(1):WA319-WA334.

doi: 10.1190/geo2022-0273.1

15. Ji M, Zhao X, Zhu W, *et al*. Noise attenuation using adaptive wavelet threshold based on CEEMD inf-X domain. *J Seism Explor*. 2023;32(2):131-153.

16. Yin X, Liu Q, Huang X, Pan Y. Development and application of a novel hybrid CEEMDAN-LWT denoising approach in microseismic monitoring. *Pure Appl Geophys*. 2022;179(9):3279-3294.

doi: 10.1007/s00024-022-03115-8

17. Zhang Y, Zhan K, Song P, Liang C, Xu R, Kong C. Improved first arrival picking of microseismic P-waves in coal mines using multi-denoising and adaptive characteristic functions. *Sci Rep*. 2025;15(1):33141.

doi: 10.1038/s41598-025-18503-y

18. Shi Y, Yang C, Wang J, Zhang Z, Meng F, Bai H. A forecasting model of ionospheric foF2 using the LSTM network based on ICEEMDAN decomposition. *IEEE Trans Geosci Remote Sens*. 2023;61:1-16.

doi: 10.1109/TGRS.2023.3336934

19. Li X, Li H, Yang Z, Zhou J, Li H, Bu J. Radiation signal denoising method of loaded coal-rock based on ICEEMDAN-PCK-means-IP. *IEEE Sens J*. 2023;23(19):23103-23118.

doi: 10.1109/JSEN.2023.3306932

20. Nguyen VT, Nguyen MT. Advanced noise reduction for in-cylinder combustion pressure data using ICEEMDAN and optimal wavelet selection. *IEEE Access*. 2025;13:62029-62044.

doi: 10.1109/ACCESS.2025.3558428

21. Esangbedo MO, Xue J, Bai S, Esangbedo CO. Relaxed rank order centroid weighting MCDM method with improved grey relational analysis for subcontractor selection: Photothermal power station construction. *IEEE Trans Eng Manage*. 2024;71:3044-3061.

doi: 10.1109/TEM.2022.3204629

22. He J, Wang H, Wang J, Zuo C, Wang Z, Zhao X. Decision-making method for multiobjective demagnetization strategies of underwater vehicles based on improved gray relation analysis. *IEEE Trans Instrument Measure*. 2024;73:1-13.

doi: 10.1109/TIM.2023.3341120

23. Tan R, Zhang W, Chen S. Decision-making method based on grey relation analysis and trapezoidal fuzzy neutrosophic numbers under double incomplete information and its application in typhoon disaster assessment. *IEEE Access*. 2020;8:3606-3628.

doi: 10.1109/ACCESS.2019.2962330

24. Fu J, Fan R, Cao J, Zhang X, Shi S. Seismic impedance inversion using a joint deep learning model based on convolutional neural network and transformer. *IEEE J Select Topics Appl Earth Observ Remote Sens*. 2023;16:8913-8922.

doi: 10.1109/JSTARS.2023.3318078

25. Yang C, Guan Y, Liu T, Zhang B, Li W. MMI-ML: Maximize mutual information between different views for few-shot remote sensing image classification. *IEEE Geosci Remote Sens Lett*. 2024;21:1-5.

doi: 10.1109/LGRS.2024.3388463

26. Pang C, Ding W, Cheng C, *et al*. Research on seismic discrimination based on particle swarm optimization generalized regression neural network and HHT sample entropy. *Prog Geophys*. 2022;37(4):1457-1463.

doi: 10.6038/pg2022FF0438

27. Mendo L. Estimation of a probability with guaranteed normalized mean absolute error. *IEEE Commun Lett*. 2009;13(11):817-819.

doi: 10.1109/LCOMM.2009.091128

28. Liu N, Wu L, Wang J, Wu H, Gao J, Wang D. Seismic data reconstruction via wavelet-based residual deep learning. *IEEE Trans Geosci Remote Sens*. 2022;60:1-13.

doi: 10.1109/TGRS.2022.3152984

29. Song J, Ntibahanana M, Luemba M, Tondozi K, Imani G. Ensemble deep learning-based porosity inversion from seismic attributes. *IEEE Access*. 2023;11:8761-8772.

    doi: 10.1109/ACCESS.2023.3239688

30. Dou Y, Li K, Lv W, Li T, Xiao Y. ContrasInver: Ultra-sparse label semi-supervised regression for multidimensional seismic inversion. *IEEE Trans Geosci Remote Sens*. 2024;62:1-13.

    doi: 10.1109/TGRS.2024.3410022

31. Pourmirzaee R, Hosseini S. Development of an ANN-based technique for inversion of seismic refraction travel times. *J Environ Eng Geophys*. 2024;29(2):75-90.

    doi: 10.32389/JEEG22-044

32. Yan Y, Chen X, Li J, Guan J, Li Y, Cui S. Multiparameter shallow-seismic waveform inversion based on the Jensen-Shannon divergence. *Geophys J Int*. 2024;238(1):132-155.

    doi: 10.1093/gji/ggae143

33. Kumar DR, Samui P, Burman A, Kumar S. Seismically induced liquefaction potential assessment by different artificial intelligence procedures. *Transport Infrast Geotechnol*. 2024;11(3):1272-1293.

    doi: 10.1007/s40515-023-00327-w

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Seismic azimuthal amplitude variation with offset response analysis in coalbed methane reservoirs with aligned fractures using Biot theory

**ZhaoJi Zhang[1], Fei Gong[1,2]*, GuanGui Zou[1,2]*, Qiang Guo[3], GuoWei Zhu[1,2], and Hao Li[1]**

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

[2]State Key Laboratory for Fine Exploration and Intelligent Development of Coal Resources, College of Geoscience and Surveying Engineering, China University of Mining and Technology–Beijing, Beijing, China

[3]Department of Geophysics, School of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu, China

***Corresponding authors:**
Fei Gong
(cupgongfei@163.com)
GuanGui Zou
(cumtzgg@foxmail.com)

**Citation:** Zhang ZJ, Gong F, Zou GG, Guo Q, Zhu GW, Li H. Seismic azimuthal amplitude variation with offset response analysis in coalbed methane reservoirs with aligned fractures using Biot theory. *J Seismic Explor.* 2026;35(2):025470118. doi: 10.36922/JSE025470118

**Received:** November 22, 2025

**Revised:** February 3, 2026

**Accepted:** February 10, 2026

**Published online:** March 4, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Abstract

The development of coalbed methane (CBM) relies on high-precision reservoir prediction and lithological inversion. Seismic amplitude variation with offset (AVO) technology is an important tool for fine-scale reservoir characterization. However, the seismic AVO response of CBM reservoirs is complex and is affected by seismic rock physics parameters at different scales. Microscopically, aligned fractures in CBM reservoirs produce complex anisotropy due to formation inclination. At the macroscopic scale, the thickness of CBM reservoirs within seismic frequency bands is comparable to the seismic wavelength and should therefore be treated as a layered medium. In addition, pore fluid significantly affects seismic wave propagation. Consequently, determining the azimuthal AVO response of CBM reservoirs in relation to seismic rock physics parameters at different scales can support high-precision reservoir prediction and lithological inversion. In this study, the primary–primary wave reflection coefficient for a two-phase layered medium was derived using Biot's theory. Using this model, the response characteristics of the reflection coefficients with respect to seismic azimuth, aligned fracture parameters, reservoir thickness, and seismic main frequency were analyzed. A rotated staggered–grid finite–difference algorithm was employed to simulate wavefield characteristics separately for coal seams and surrounding strata. Seismic attributes were then used to characterize the seismic AVO response. The Z-direction seismic amplitude attributes and reflection coefficients showed similar trends in their responses to seismic rock-physical parameters. This study contributes to establishing a more precise seismic AVO response framework to support CBM reservoir prediction and high-quality lithological inversion.

***Keywords:*** Azimuth; Amplitude variation with offset; Tilted transversely isotropic anisotropy; Fracture; Modeling

## 1. Introduction

Coalbed methane (CBM) is an unconventional natural gas with significant development potential.[1,2] The development of CBM depends on the availability of accurate reservoir predictions. Seismic amplitude variation with offset (AVO) response is a significant tool for reservoir prediction and characterization, as it contains information on underground lithology, pore space, fluid properties, and numerous other parameters.[3] Nevertheless, the seismic AVO response of CBM reservoirs is influenced by many factors, and the governing mechanisms remain unclear. At present, this limits its application as a reliable basis for lithological inversion and accurate CBM reservoir prediction.

Seismic AVO response characteristics are important indicators for reservoir prediction.[4,5] In studies of reservoir AVO responses, Xu *et al.*[6] used physical modeling to predict the AVO response laws of reservoirs of different sizes. Xiao *et al.*[7] investigated the AVO response characteristics of natural gas hydrates with varying properties and established a corresponding rock physics model. Gavin and Lumley[8] studied azimuthal AVO response characteristics in anisotropic media. Furthermore, Liu *et al.*[9] simulated the seismic response of horizontally fractured, thin, interbedded gas reservoirs using physical modeling techniques and analyzed the frequency-dependent AVO response under the combined effects of thin-layering and fractures. Xi and Yin[10] studied the seismic reflection laws of various sandstone formations and coal seams, summarizing the seismic response of common coal-bearing reservoir lithology combinations. Kumar *et al.*[11] developed a frequency-variant AVO (FAVO) model and combined it with logging data, demonstrating that reservoir stacking patterns are the primary factor controlling FAVO tuning response, followed by pore fluids. Following this, Ahmed *et al.*[12] developed a FAVO inversion method incorporating the quality factor, while Ouyang *et al.*[13] compared different FAVO inversion methods and demonstrated the superiority of P-wave dispersion for reservoir prediction, while Tian *et al.*[14] proposed a deep neural network-based method for predicting reservoir fluid parameters. The above studies show that research on seismic AVO responses mainly focuses on two aspects. On the one hand, it focuses on microparameters in unconventional reservoirs, analyzing the seismic AVO responses to factors such as reservoir microstructure, pore parameters, and pore fluids.[5,7,11-19] On the other hand, it investigates seismic AVO responses related to reservoir tectonic characteristics, such as reservoir thickness and seismic frequency.[6,8,10,20-22]

Exploration of CBM requires increasingly accurate seismic exploration to portray reservoir lithology, fracture patterns, fluid types, and other microscopic parameters.[23] However, macroscopic factors such as reservoir thickness, azimuth, and seismic frequency can introduce uncertainties into reservoir prediction. Several issues remain unresolved in the above two areas of seismic AVO response research. In research focusing on microscopic factors, the seismic AVO response associated with fracture inclination remains unclear. Previous studies have shown that porosity fracture inclination angle significantly affects reservoir anisotropy and AVO response characteristics of reservoir rock.[24,25] Large-dip coal seams are widely distributed in China.[26,27] However, many current studies equate coal seams with high-angle fractures to horizontal transverse isotropic media. There are significant differences in AVO responses between these two, yet the seismic wavefield response in three-dimensional space with respect to fracture inclination angle has not been sufficiently considered.[28-30] In the study of the seismic azimuthal AVO response to macroscopic factors, the seismic AVO response of coal seams containing inclined fractures remains unclear. The thickness of unconventional natural gas reservoirs is often comparable to or smaller than the seismic wavelength, meaning that such reservoirs cannot be treated as a homogeneous infinite medium, but should instead be considered as a layered medium. Seismic wave reflections in a layered medium lead to interference effects that influence AVO response characteristics. Although the effects of reservoir thickness and seismic frequency are significant, the combined influence of these factors on the seismic AVO response of CBM reservoirs has not been adequately considered.

Rock physics can be used to quantitatively describe how rock microstructure, fracture morphology, and related parameters can affect elastic properties. Rock physics theory can be used to establish relationships among reservoir physical properties, microstructure, and seismic AVO response. Huang *et al.*[31] modeled the rock physics of CBM reservoirs and analyzed the relationship between intercept and gradient responses and CBM content. Wu *et al.*[32] quantified the effects of horizontal fractures on elasticity and various anisotropies in coal seams using rock physics modeling. Gong *et al.*[33] developed an anisotropic, multi-porosity rock physics model to predict the effects of different pore types on the elasticity and anisotropy of CBM reservoirs. These studies mainly focused on mapping the relationships between coalbed microstructure and elasticity and anisotropy, but several limitations remain. First, the seismic response of pore fluid under frequency-dependent conditions is often neglected, as the Gassmann equation is commonly applied to incorporate pore fluid without accounting for frequency effects in the modeling process.[34] Reflection and transmission of seismic waves

in the formation, based on the pore medium assumption, have not received much attention, which restricts accurate characterization of the seismic AVO response of coal seams. Second, fracture morphology is insufficiently accounted for in modeling, which typically treats fractures as vertical or horizontal distributions without considering their inclination angles. Furthermore, rock physics modeling primarily focuses on mapping the relationships between rock microstructures and elastic features, yet it does not explicitly depict the AVO response characteristics of seismic waves between strata. Seismic AVO response analysis combines rock physics modeling with coal stratigraphic features to determine the seismic AVO response characteristics under the influence of multiscale factors. This approach accurately portrays CBM reservoirs and enables high-precision lithological inversion.

Existing research has predominantly focused on the microphysical properties of reservoirs. Owing to the complexity of CBM reservoirs, studies that explicitly address the coupled effects of anisotropic fractures, seismic frequency dispersion, and thin-bed tuning on the azimuthal AVO response remain scarce. This gap hinders the accurate interpretation of seismic data for CBM reservoir characterization. In this study, the formulas for calculating the reflection coefficients of P-waves in a two-phase layered medium are first derived using Biot's theory. Combined with rock physics modeling, changes in the reflection coefficient in CBM reservoirs under multi-scale factors, such as fracture morphology, frequency, and layer thickness, are predicted. Seismic waveform response characteristics, including azimuth, oriented fracture porosity, fracture inclination, and fluid bulk modulus, were simulated using a finite-difference algorithm with a rotated staggered grid. The response laws of the seismic reflection coefficients and the seismic responses of CBM reservoirs with respect to multi-scale parameters, such as fracture morphology, frequency, and layer thickness, are summarized. This study provides a foundation for predicting CBM reservoirs and for high-precision lithological inversion by combining and guiding the seismic response law in complex media.

## 2. Methods

### 2.1. Coalbed methane reservoir modeling with aligned fracture inclination angle and azimuth

The pore structure of CBM reservoirs is highly complex. It consists of two types of pores: matrix pores and fractures.[35] Matrix pores and microscopic fractures have poor connectivity. These are the primary adsorption areas for CBM, whereas the macroscopic fractures have better connectivity and act as the main seepage channels for coalbed water and CBM. Aligned fractures are the main source of anisotropy in CBM reservoirs.[33,36] Large-dip coal seams are widely distributed in China,[26,27] and fractures with different inclination angles can affect the azimuthal AVO response of CBM reservoirs. A rock-physics model of CBM reservoirs, incorporating multiple fracture inclination angles and azimuths, was developed to study the three-dimensional seismic AVO response of complex fracture systems. The Voigt–Reuss–Hill theory was used to construct the coal matrix, which consists of organic matter (OM) and ash. To restore the double-pore characteristics and anisotropy of the CBM reservoir, the self-consistent approximation theory was used to calculate the matrix pore, and the Schoenberg theory was used to calculate the aligned fractures. As shown in Figure 1, the bond transform rotates the medium around the x- and z-axes by a given angle. The fracture inclination angle is the angle between the fracture plane and the XOY plane, and the fracture azimuth is the angle with the Y-axis. The effect of pore fluid in coal cannot be ignored. Dispersion and attenuation in elastic media under the action of seismic waves mainly come from the pore fluid.[37] Within the seismic frequency band, the fluid relaxes within the pore space, and the pore pressure increment induced by seismic waves is equilibrated throughout the pore space.[38] Building on this, Biot[39] proposed a theoretical model of elastic wave dispersion attenuation on a macroscopic scale.[39] This model aims to restore the seismic response characteristics of CBM reservoirs under the influence of dispersion attenuation in seismic frequency bands. A rock-physics model of a CBM reservoir was constructed by incorporating a fluid phase into Biot's theory. Assuming the top and bottom plates of the coal bed are isotropic mudstone layers, the specific modeling process is shown in Figure 2.

### 2.2. Effect of aligned fractures on reflection coefficients of two-phase layered media

#### *2.2.1. Primary-primary wave reflection coefficient in two-phase layered media*

The reflection coefficient bridges the gap between the formation's elastic properties and its seismic response. It can characterize the strength of reflected wave amplitudes and the AVO response of reservoirs. During seismic propagation in CBM reservoirs, the wavelengths of seismic wavelets are similar to the thickness of the coal seam. In this case, the coal seam can be considered a laminar medium. The P–P reflection coefficient (RPP) of a laminar medium differs somewhat from that of an infinite homogeneous medium in a half-space. Previous researchers have used the propagation matrix theory in layered media to study thin layers and to derive their equivalent reflection and transmission coefficients.[40-43] The RPP formulation for

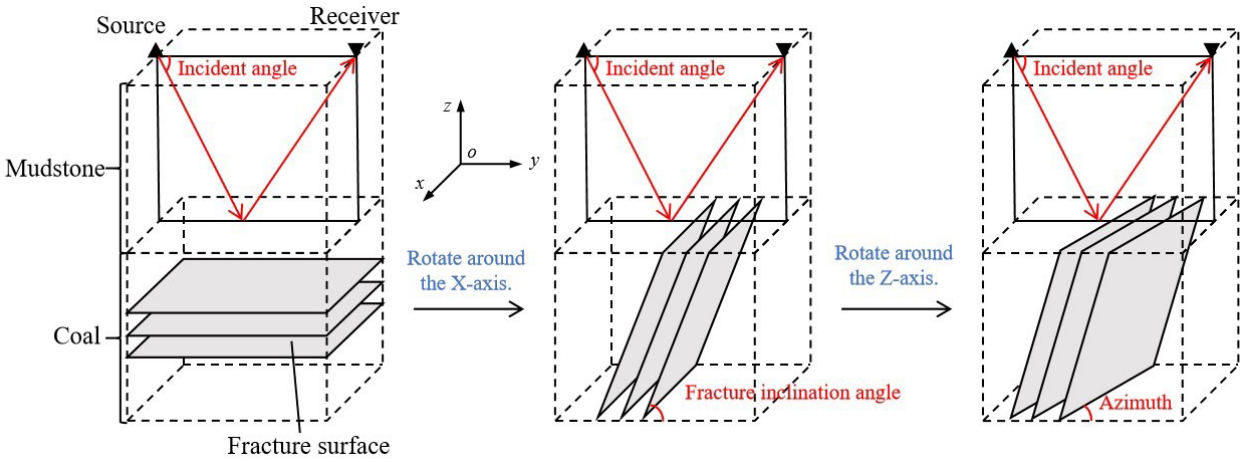



**Figure 1.** The coordinate transformations of the fracture inclination angle, incident angle, and azimuth

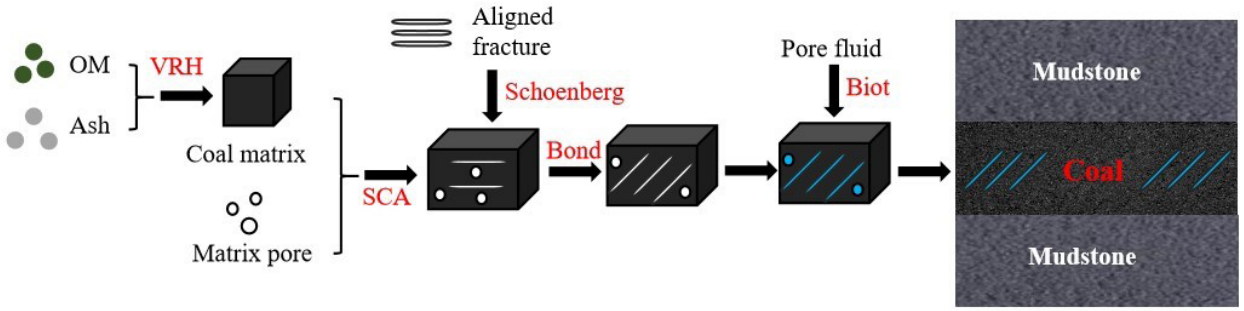



**Figure 2.** Rock physics model workflow
Abbreviations: OM: Organic matter; SCA: Self-consistent approximation; VRH: Voigt–Reuss–Hill.

two-phase layered media has been developed to enable more detailed characterization of CBM reservoirs within seismic frequency bands, supporting high-precision lithology prediction.

Forward modeling studies of thin layers typically employ the following two approaches:[44]

(i) Establishing reflection and transmission relationships for thin layers using time delay under arbitrary incidence angles.[45,46] This method is not suitable for investigating AVO characteristics of single thin layers.[47]

(ii) Employing the propagation matrix theory for layered media to investigate thin-layer problems.[40-43] This approach further simplifies the P-wave and S-wave reflection theory for multi-layer media under arbitrary incidence angles to derive the P-wave reflection coefficient formula for a single thin layer.

We employed the theory of layered medium propagation to address thin-layer reflection phenomena. The generalized Hooke's law in two-phase media is shown in **Equation 1**:

$$\begin{bmatrix}\sigma \\ s\end{bmatrix}=\begin{bmatrix}C & Q \\ Q & R\end{bmatrix}\begin{bmatrix}e \\ \varepsilon\end{bmatrix} \tag{1}$$

where $C$ is the solid phase elastic stiffness matrix, $\sigma$ is the force tensor in the solid phase, and $s$ is the effective stress in the fluid phase. $e$ and ε are the strains in the solid and fluid phase, respectively. $Q$ is the fluid-solid coupling term, and the method for the value of $Q$ in anisotropic media is derived from Dai *et al.*[48] $R$ is the fluid pressure coefficient, which is determined by the bulk modulus of the fluid and the porosity of the medium,[49] as shown in **Equation 2**:

$$R=\frac{\phi^2}{(1-\phi-\beta)C_r+\phi/K_f} \tag{2}$$

where $\phi$ is the porosity, $\beta$ is the ratio of the compression coefficients of the two-phase media particles and the skeleton; $C_r$ is the compression coefficient of the two-phase media particles. and $K_f$ is the fluid bulk modulus.

As shown in Figure 3, it is assumed that a P1 wave incident from an infinite homogeneous medium 1 passes through a series of thin layers (2 to $m$−1) of medium and

finally enters the lower medium *m*, where it is reflected from the n-th layer. In the n-th layer, there are six waves: the incident P1 wave, the reflected P1 wave, the incident S wave, the reflected S wave, the incident P2 wave, and the reflected P2 wave. Their wave functions can be set as follows:

(i) Incident P1 wave (**Equation 3**):

$$\begin{cases} \varphi_{s1} = A_1 \exp\left[ I\left(kx_1 x + kz_1 z - wt\right)\right] \\ \varphi_{f1} = A_{f1} \exp\left[ I\left(kx_1 x + kz_1 z - wt\right)\right] \end{cases} \tag{3}$$

(ii) Incident P2 wave (**Equation 4**):

$$\begin{cases} \varphi_{s2} = A_2 \exp\left[ I\left(kx_2 x + kz_2 z - wt\right)\right] \\ \varphi_{f2} = A_{f2} \exp\left[ I\left(kx_2 x + kz_2 z - wt\right)\right] \end{cases} \tag{4}$$

(iii) Incident S wave (**Equation 5**):

$$\begin{cases} \psi_{s3} = A_3 \exp\left[ I\left(kx_3 x + kz_3 z - wt\right)\right] \\ \psi_{f3} = A_{f3} \exp\left[ I\left(kx_3 x + kz_3 z - wt\right)\right] \end{cases} \tag{5}$$

(iv) Reflected P1 wave (**Equation 6**):

$$\begin{cases} \varphi_{s4} = A_4 \exp\left[ I\left(kx_4 x + kz_4 z - wt\right)\right] \\ \varphi_{f4} = A_{f4} \exp\left[ I\left(kx_4 x + kz_4 z - wt\right)\right] \end{cases} \tag{6}$$

(v) Reflected P2 wave (**Equation 7**):

$$\begin{cases} \varphi_{s5} = A_5 \exp\left[ I\left(kx_5 x + kz_5 z - wt\right)\right] \\ \varphi_{f5} = A_{f5} \exp\left[ I\left(kx_5 x + kz_5 z - wt\right)\right] \end{cases} \tag{7}$$

(vi) Reflected S wave (**Equation 8**)

$$\begin{cases} \psi_{s6} = A_6 \exp\left[ I\left(kx_6 x + kz_6 z - wt\right)\right] \\ \psi_{f6} = A_{f6} \exp\left[ I\left(kx_6 x + kz_6 z - wt\right)\right] \end{cases} \tag{8}$$

where $kx_i = \frac{\sin\theta_i}{v_i} w$, $kz_i = \frac{\cos\theta_i}{v_i} w$, in which θ is the angle of incidence or reflection of the wave. $\varphi_s$ and $\varphi_f$ are the displacements in the solid and fluid phase, respectively. $A_i$ and $A_{fi}$ are the amplitudes of the solid and fluid parts, respectively, and let $A_{fi} = b_i A_i$. *v* is the velocity of the wave, *w* is the circular frequency, and *I* is the imaginary unit.

From Snell's law, we can obtain **Equation 9**:

$$kx_i = \frac{\sin\theta_1}{v_1} w = \frac{\sin\theta_2}{v_2} w = \cdots = \frac{\sin\theta_6}{v_6} w = \delta \tag{9}$$

The displacement component of the solid phase medium is set to be $u = [u_x, u_y, u_z]^T$ and the displacement component of the fluid phase medium be $U = [U_x, U_y, U_z]^T$. The displacement function can be expressed as **Equation 10**:

$$\begin{cases} u = \nabla\varphi_s + \nabla\times\psi_s \\ U = \nabla\varphi_f + \nabla\times\psi_f \end{cases} \tag{10}$$

where, $\varphi_s$ and $\psi_s$ are the bit functions of the P-wave and S-wave displacements in the solid phase, respectively. $\varphi_f$ and $\psi_f$ are bit functions of the P-wave and S-wave displacements in the fluid phase, and the Lorentz condition is satisfied between them. The displacements and displacement bits in the medium are shown in **Equation 11**:

$$\begin{cases} \varphi_s = \varphi_{s1} + \varphi_{s2} + \varphi_{s4} + \varphi_{s5} \\ \varphi_f = \varphi_{f1} + \varphi_{f2} + \varphi_{f4} + \varphi_{f5} \\ \psi_s = \psi_{s3} + \psi_{s5} \\ \psi_f = \psi_{f3} + \psi_{f5} \end{cases} \tag{11}$$

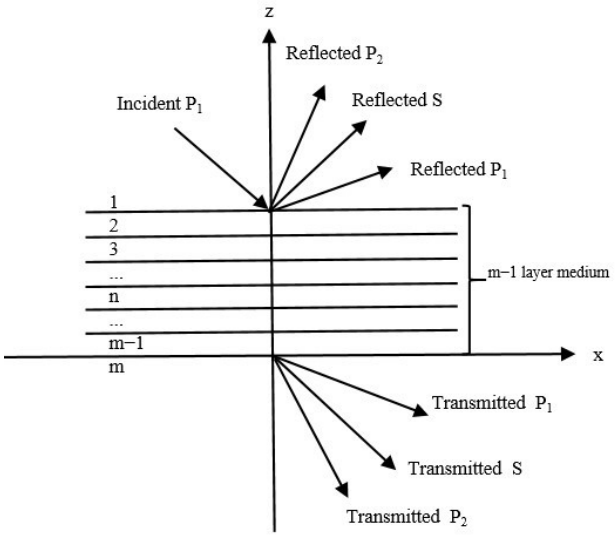



**Figure 3.** The propagation of seismic waves in layered media

Six boundary continuous conditions exist on the boundary plane of a two-phase medium:[49,50]

(i) Solid tangential displacement $u_x$ continuum condition (**Equation 12**):

$$u_x^{(1)} = u_x^{(2)} \tag{12}$$

(ii) Solid normal displacement $u_z$ continuum condition (**Equation 13**):

$$u_z^{(1)} = u_z^{(2)} \tag{13}$$

(iii) Continuous conditions of fluid mass *M* at the

interface (**Equation 14**):

$$M^{(1)} = M^{(2)} \tag{14}$$

where $M = \rho_f \frac{\partial \varphi_f}{\partial z} + \rho_f \frac{\partial \psi_f}{\partial x}$, and $\rho_f$ is the fluid density.

(iv) Total normal stress $\sigma_{zz} + s$ continuity condition (**Equation 15**):

$$\sigma_{zz}^{(1)} + s^{(1)} = \sigma_{zz}^{(2)} + s^{(2)} \tag{15}$$

(v) Solid tangential stress $\tau$ continuum condition (**Equation 16**):

$$\tau^{(1)} = \tau^{(2)} \tag{16}$$

(vi) Fluid pressure $P$ continuity conditions (**Equation 17**):

$$P^{(1)} = P^{(2)} \tag{17}$$

where $s = -\phi P$. The superscripts 1 and 2 represent physical quantities in different upper and lower media, respectively. A recursive formula is established to determine the boundary conditions at the top and bottom interfaces of the thin-layer medium. The detailed derivation process is provided in the Appendix.

$$\begin{bmatrix} u_x^{(n)} \\ u_z^{(n)} \\ M^{(n)} \\ \sigma_{zz}^{(n)} + s^{(n)} \\ \tau_{xz}^{(n)} \\ -P^{(n)} \end{bmatrix} = [H_{m-1}]_{z=Z(m-1)} \cdot d_{m-1} \cdots d_3 \cdot [H_2]^{-1}_{z=Z(1)} \cdot \begin{bmatrix} u_x^{(1)} \\ u_z^{(1)} \\ M^{(1)} \\ \sigma_{zz}^{(1)} + s^{(1)} \\ \tau_{xz}^{(1)} \\ -P^{(1)} \end{bmatrix} \tag{18}$$

where $d_n = [H_n]^{-1}_{z=Z(n-1)} \cdot [H_{n-1}]_{z=Z(n-1)}$. $Z(n)$ is the coordinate of the n-th layer bottom on the z-axis. By substituting the wave function expressions for layer 1 and layer $m$ into **Equation 18** and combining with the method for deriving RPPs in infinite homogeneous half-space media published in past studies,[51-53] one can obtain the RPP for two-phase layered media.

### *2.2.2. Characterization of reflection coefficient response for multi-scale parameters*

Based on the formula for RPP in a two-phase layered medium, this study discusses the response of RPP to seismic rock physics parameters at different scales. At the micro-scale of the coal seam, parameters affecting fracture morphology include fracture porosity and inclination. Conversely, at the seismic macro-scale, factors such as seismic azimuth, coal seam thickness, and dominant wavelet frequency are important. This study aims to investigate the impact of multi-scale parameters on the RPP of a single coal seam. The top and bottom plates of the coal seam were set as mudstone with P-wave velocity of 3,500 m/s and S-wave velocity of 2,100 m/s. The coal matrix composition was set to 90% OM (with a bulk modulus of 5 GPa and a shear modulus of 2.5 GPa) and 10% ash (with a bulk modulus of 40 GPa and a shear modulus of 20 GPa). The matrix pore porosity was set to 3%, and the matrix pore aspect ratio was set to 0.2. The fracture aspect ratio was set to 0.01. To investigate the RPP response with respect to the azimuth, incidence angle, and slit inclination angle, the slit porosity was set to 2%, the seismic wavelet frequency was set to 50 Hz, and the coal seam thickness was set to 5 m. The responses of RPP to slit inclination angles and azimuths at different incidence angles are shown in Figure 4A-D, corresponding to incidence angles of 0–30°, respectively. As shown in Figure 4, the RPP value first increases and then decreases with increasing fracture inclination angle. The position of the peak depends on the incidence angle and azimuth. The RPP variation is axisymmetric about x = 180°, and the larger the incidence angle, the more drastic the change on either side of the symmetry axis. As the fracture inclination angle and incidence angle increased, RPP gradually changed from a single peak to a double peak with azimuthal variation.

The incidence angle and azimuth were set to 0°. Figure 5 shows the response relationship between RPP and fracture porosity and fracture inclination angle, where Figure 5A-D correspond to incidence angles of 0–30°, respectively. As shown in Figure 5, the RPP initially increases and then decreases with increasing fracture inclination angle, peaking at approximately 35°. The fracture inclination angle is an important factor influencing the medium's anisotropy. The medium can be approximated as vertical transverse isotropy anisotropy at low fracture inclination angle, as horizontal transverse isotropy anisotropy at high fracture inclination angle, and at medium fracture inclination angle as tilted transverse isotropy anisotropy. The RPP values in different anisotropic media exhibited distinct patterns as fracture porosity increased. At a high fracture inclination angle, larger incident angles produced greater changes in RPP with respect to fracture porosity. At a low fracture inclination angle, the RPP response to fracture porosity depended on the incident angle: it increased with increasing fracture porosity at 0° and 10°, but decreased at 30°.

The Biot–Gassmann–Wood relationship describes the effect of mixed fluids on rock elastic properties

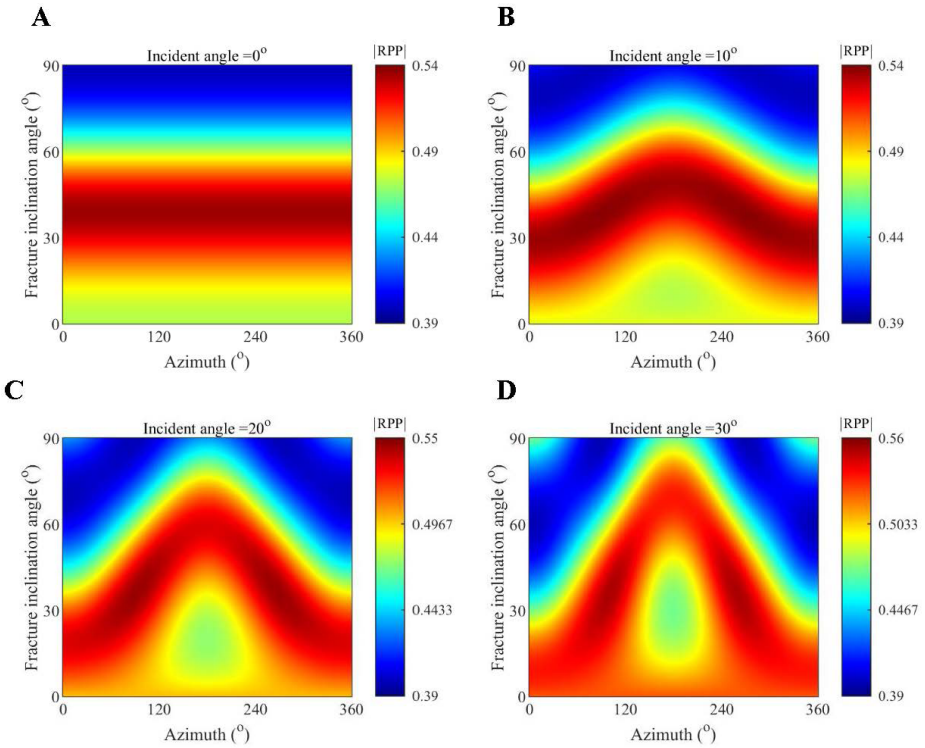



**Figure 4.** Response of the P–P reflection coefficient (RPP) with respect to fracture inclination angle and azimuth for different incident angles. (A) 0°, (B) 10°, (C) 20°, and (D) 30°.

across different frequencies.[38,54] The pore fluids in CBM reservoirs are complex and exist in adsorbed, free, and dissolved forms.[33] Different fluid storage forms and fluid properties mainly affect the bulk modulus and density of mixed pore fluids.[55] To study the response laws of the RPP with pore fluid parameters, the fracture inclination angle and azimuth of the CBM reservoir were set at 0°, and the porosity of the fracture was 2%. The response of the RPP of the CBM reservoir on the bulk modulus and density of the mixed fluid is shown in Figure 6. The RPP decreases as the bulk modulus of the pore fluid increases, while it varies little with density.

For layered media, layer thickness and seismic wavelet frequency are important parameters that affect the tuning effect. To investigate the effects of seismic wavelet frequency and coal seam thickness on the RPP, the fracture inclination angle was set to 45°, and the incidence angle and azimuth were set to 0°. The RPP was then predicted as a function of coal seam thickness and the dominant frequency of the seismic wavelet. As shown in Figure 7, when the seismic dominant frequency is low and the coal seam is thin, the seismic wavelength is much larger than the coal seam thickness, resulting in a very small RPP. As the seismic wave frequency and coal seam thickness increased, the RPP of the coal seam showed a periodic relationship with both thickness and frequency. With increasing frequency, the RPP variation with coal seam thickness becomes increasingly pronounced.

## 3. Seismic wavefield simulation

### 3.1. Numerical simulation method

The finite-difference method can simulate the wavefield characteristics of seismic waves in a medium. The resulting seismic record better reflects the reflection and refraction characteristics of seismic waves than the convolution method. High-precision finite differences can simulate the propagation laws of seismic waves in two-phase pore media and quantitatively analyze the seismic response characteristics of different fracture patterns in coal. The displacement and stress components can be expressed as **Equation 19**:

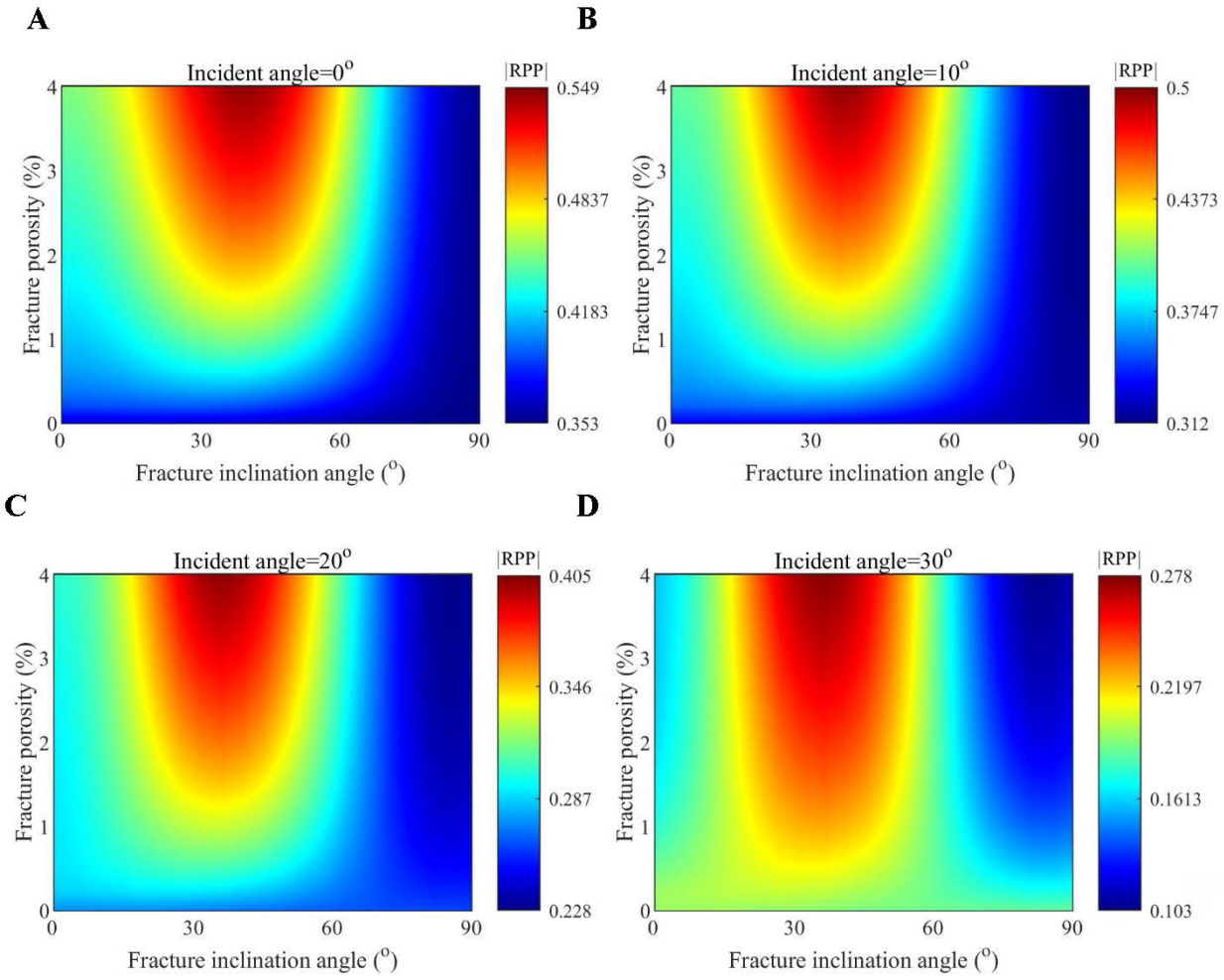



**Figure 5.** Response of the P–P reflection coefficient (RPP) with respect to fracture inclination angle and fracture porosity. (A) 0°, (B) 10°, (C) 20°, and (D) 30°.

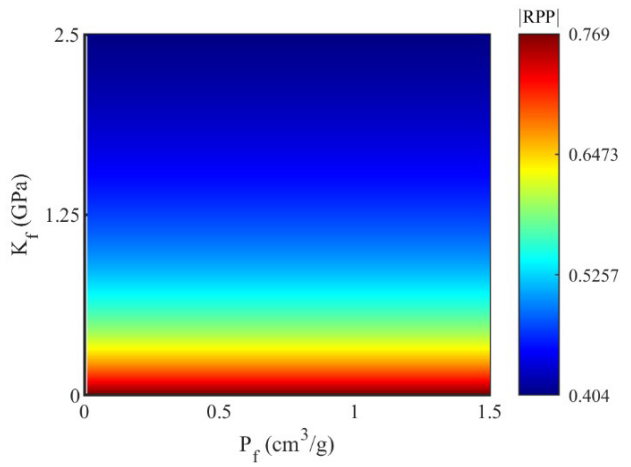



**Figure 6.** Response of the reflection coefficient with respect to pore fluid parameters. $K_f$: bulk modulus of the pore fluid; $P_f$: Pore fluid pressure; RPP: P–P reflection coefficient.

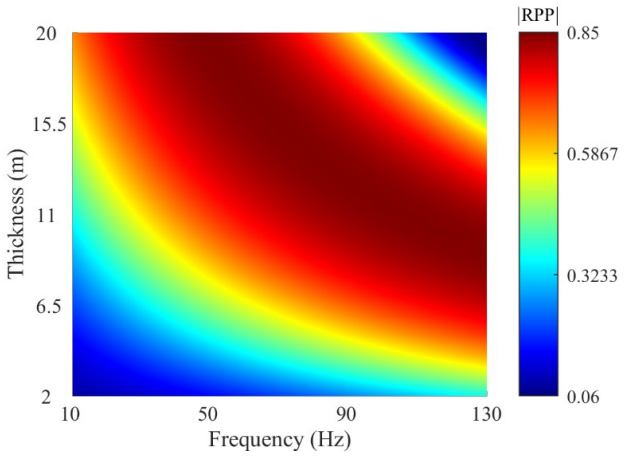



**Figure 7.** P–P reflection coefficient (RPP) on the response of coal seam thickness to seismic wave main frequency

$$\begin{cases}
\frac{\partial u_x}{\partial t}=(D_2+D_3)b_{11}(u_x-U_x)-D_3\left(\frac{\partial \tau_{xx}}{\partial x}+\frac{\partial \tau_{xz}}{\partial z}\right)+D_2\frac{\partial s}{\partial x}\\
\frac{\partial u_y}{\partial t}=(D_2+D_3)b_{22}(u_y-U_y)-D_3\left(\frac{\partial \tau_{xy}}{\partial x}+\frac{\partial \tau_{yz}}{\partial z}\right)\\
\frac{\partial u_z}{\partial t}=(D_2+D_3)b_{33}(u_z-U_z)-D_3\left(\frac{\partial \tau_{xz}}{\partial x}+\frac{\partial \tau_{zz}}{\partial z}\right)+D_2\frac{\partial s}{\partial z}\\
\frac{\partial U_x}{\partial t}=-(D_1+D_2)b_{11}(u_z-U_z)+D_2\left(\frac{\partial \tau_{xx}}{\partial x}+\frac{\partial \tau_{xz}}{\partial z}\right)-D_1\frac{\partial s}{\partial x}\\
\frac{\partial U_y}{\partial t}=-(D_1+D_2)b_{22}(u_z-U_z)+D_2\left(\frac{\partial \tau_{xy}}{\partial x}+\frac{\partial \tau_{yz}}{\partial z}\right)\\
\frac{\partial U_z}{\partial t}=-(D_1+D_2)b_{33}(u_z-U_z)+D_2\left(\frac{\partial \tau_{xz}}{\partial x}+\frac{\partial \tau_{zz}}{\partial z}\right)-D_1\frac{\partial s}{\partial z}\\
\frac{\partial \sigma_{xx}}{\partial t}=C_{11}\frac{\partial u_x}{\partial x}+C_{13}\frac{\partial u_z}{\partial z}+C_{14}\frac{\partial u_y}{\partial z}+C_{15}(\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z})+C_{16}\frac{\partial u_y}{\partial x}+Q_1\left(\frac{\partial U_x}{\partial x}+\frac{\partial U_z}{\partial z}\right)\\
\frac{\partial \sigma_{zz}}{\partial t}=C_{13}\frac{\partial u_x}{\partial x}+C_{33}\frac{\partial u_z}{\partial z}+C_{34}\frac{\partial u_y}{\partial z}+C_{35}(\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z})+C_{36}\frac{\partial u_y}{\partial x}+Q_3\left(\frac{\partial U_x}{\partial x}+\frac{\partial U_z}{\partial z}\right)\\
\frac{\partial \tau_{yz}}{\partial t}=C_{14}\frac{\partial u_x}{\partial x}+C_{34}\frac{\partial u_z}{\partial z}+C_{44}\frac{\partial u_y}{\partial z}+C_{45}(\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z})+C_{46}\frac{\partial u_y}{\partial x}\\
\frac{\partial \tau_{xz}}{\partial t}=C_{15}\frac{\partial u_x}{\partial x}+C_{35}\frac{\partial u_z}{\partial z}+C_{45}\frac{\partial u_y}{\partial z}+C_{55}(\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z})+C_{56}\frac{\partial u_y}{\partial x}\\
\frac{\partial \tau_{xy}}{\partial t}=C_{16}\frac{\partial u_x}{\partial x}+C_{36}\frac{\partial u_z}{\partial z}+C_{46}\frac{\partial u_y}{\partial z}+C_{56}(\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z})+C_{66}\frac{\partial u_y}{\partial x}\\
\frac{\partial s}{\partial t}=Q_1\frac{\partial u_x}{\partial x}+Q_3\frac{\partial u_z}{\partial z}+R\left(\frac{\partial U_x}{\partial x}+\frac{\partial U_z}{\partial z}\right)
\end{cases} \tag{19}$$

where

$$\begin{cases}
D_1=\frac{\rho_{11}}{(\rho_{12}^2-\rho_{11}\rho_{22})}\\
D_2=\frac{\rho_{12}}{(\rho_{12}^2-\rho_{11}\rho_{22})}\\
D_3=\frac{\rho_{22}}{(\rho_{12}^2-\rho_{11}\rho_{22})}
\end{cases} \tag{20}$$

and $\rho_{11}$ and $\rho_{22}$ denote the effective mass parameters in the solid and fluid phase, respectively. $\rho_{12}$ is the mass parameter of the relative motion of the pore fluid. $b_{ij}$ is the dissipation coefficient of the fluid moving relative to the skeleton.

The rotated staggered-grid finite-difference scheme was first invented, described, and applied by Saenger *et al.*[56] This technique is an extension of the traditional staggered-grid method. In models with density and velocity anomalies, the stability of the traditional finite-difference method is often insufficient.[56] The rotated staggered-grid method is used in anisotropic media to reduce errors caused by interpolation and to enhance the stability of numerical simulation.[57] Stability conditions for finite-difference calculations on rotated staggered grids were provided by Saenger *et al.*[56], as shown in **Equation 21**:

$$\frac{\Delta t V_{\max}}{\Delta h}\leqslant 1/(\sum_{k=1}^{n}|c_k|) \tag{21}$$

where $\Delta t$ is the time step, $V_{\max}$ is the maximum phase velocity, $\Delta h$ is the spatial step, and $c_k$ is the staggered grid spatial difference coefficient.

To simulate the half-space medium more accurately, the absorbing boundary condition is an important issue in the orthogonal process. The convolutional perfectly matched layer (CPML) technique was developed based on perfectly matched layers, which avoids the complexity of the conventional perfectly matched layer (PML).[58,59] CPML has adapted the form of the extended function $s_x$ as follows:[59]

$$s_x=k_x+\frac{d_x}{a_x+iw} \tag{22}$$

where $a_x\geq 0, k_x\geq 0$. When $a_x=0, k_x=1$, it is the conventional PML standard form. The CPML absorbing boundary form is derived by convolving the spatial derivatives in the time domain.

### 3.2. Characterization of wavefields in coal

As a two-phase anisotropy medium, the propagation of seismic waves in coal seams is extremely complex. The fracture inclination angle was set to 45°, the azimuth to 0°, the model size to 200 m × 200 m, and the grid spacing to 1 m. A Ricker wavelet with horizontal polarization in the x direction of the solid phase was excited at the center of the model, with a main frequency of 50 Hz. The wavefield snapshot results were obtained at t = 35 ms. Figure 8 shows the seismic wavefield result in the coal seam. As shown in Figure 8, the medium contains both fast and slow P- and S-waves. The slow P-wave has the smallest velocity. In the Biot theory, the slow P-wave originates from the relative motion of the solid phase and the fluid. It has a smaller velocity and stronger attenuation, and is essentially disregarded in seismic exploration. S-wave splitting originates from the difference between the S-wave polarization direction and the fracture plane. The polarization direction of fast S-waves is parallel to the fracture plane, whereas that of slow S-waves is perpendicular to it. Therefore, the difference between fast and slow S-waves originates from the rock's anisotropy. The difference between fast and slow P-waves originates from the interaction between the solid skeletal phase and the fluid phase of a two-phase pore medium.[36] The results

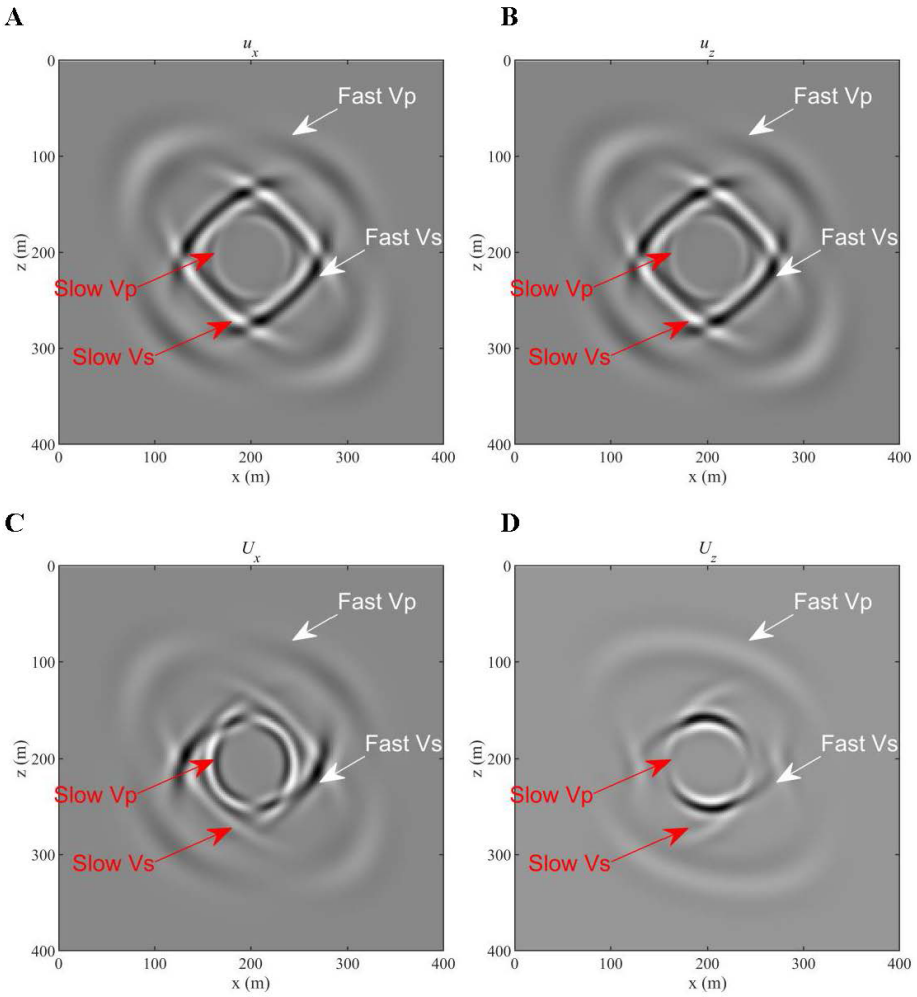



**Figure 8.** Wavefield snapshot of two-phase homogeneous TTI anisotropic media. (A) Solid phase x component, (B) solid phase z component, (C) fluid phase x component, (D) fluid phase z component.
Abbreviations: TTI: Tilted transversely isotropic; Vp: P-wave velocity; Vs: S-wave velocity.

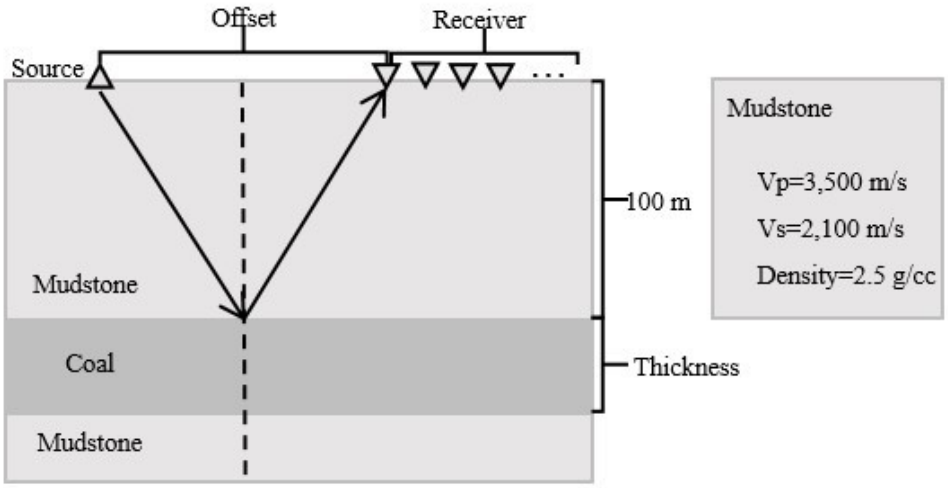



**Figure 9.** Coal-bearing strata model
Abbreviations: Vp: P-wave velocity; Vs: S-wave velocity.

demonstrate the effectiveness of the method in simulating seismic wave propagation characteristics in two-phase anisotropic media.

### 3.3. Characterization of seismic response of coalbed methane reservoirs

Seismic records in seismic exploration are mainly derived from reflected P-waves. To study the response of different seismic rock physics parameters in CBM reservoir seismic records, a seismic orthorectification model of CBM reservoirs was established, as shown in Figure 9. The coal seam top and bottom plates were set to be mudstone (P-wave velocity 3,500 m/s, S-wave velocity 2,100 m/s). The composition of the coal matrix was set to 90% OM and 10% ash, the matrix pore porosity to 3%, matrix pore asp to 0.2, fracture asp to 0.01, fracture porosity to 1%, and the pore medium is saturated with water. The seismic wavelet frequency was 50 Hz, and the thickness of the coal seam was 5 m. Full-wavefield finite-difference simulations bridged the gap between theoretical reflectivity analysis and practical seismic interpretation. They validated whether the response trends predicted by the idealized plane-wave RPP formula hold in more realistic physical scenarios, including complex wave propagation, spherical

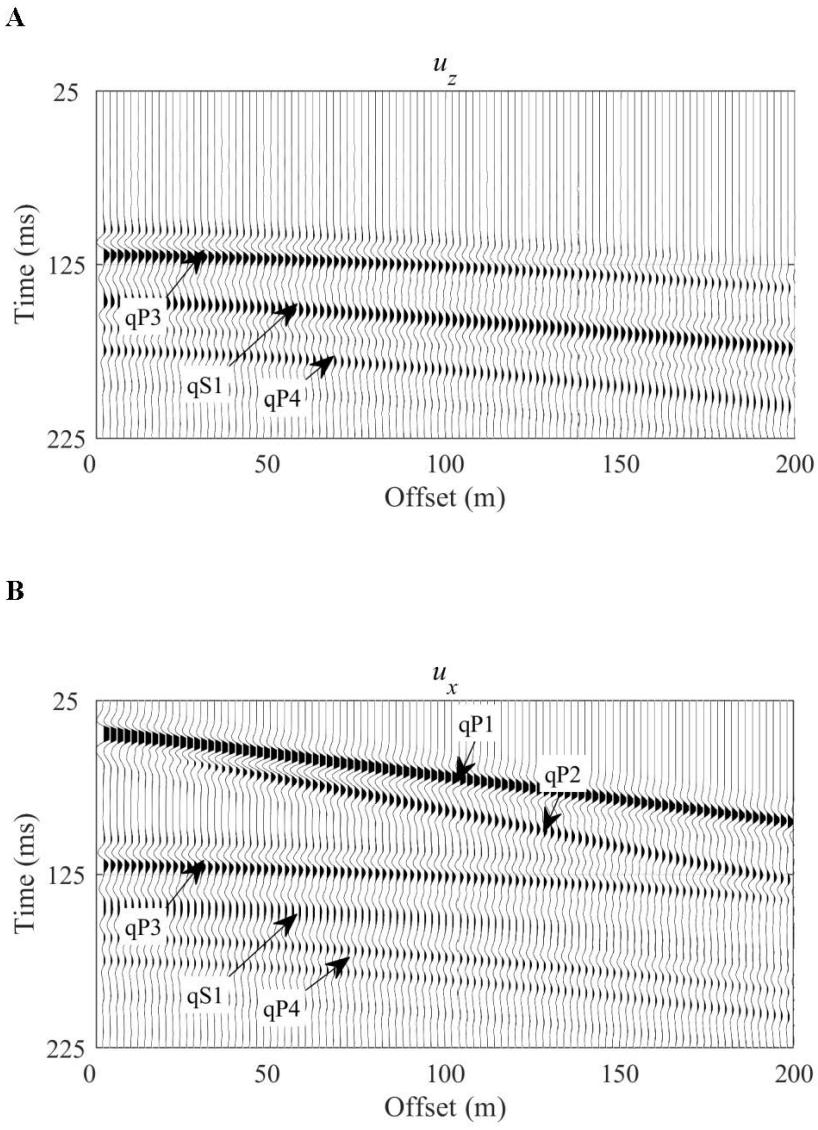



**Figure 10.** Seismic response of coalbed methane reservoirs. (A) Displacement in the z-direction. (B) x-direction displacement.

expansion, and full interference.

With the fracture inclination angle and azimuth set to 0°, Figure 10A and B shows the amplitudes $u_z$ and $u_x$ in the z and x directions, respectively, where qP1 and qP2 are the direct fast and slow P-waves, respectively. The qP3, qS1, and qP4 waves are the reflected fast P-, S-, and slow P-waves, respectively. There is a direct wave in the x direction, but no direct wave in the z direction because the line between the source and the wave detection point is parallel to the x-axis. The amplitude component of the direct wave in the z direction was zero. At near offsets, two P-waves overlapped. As the offset distance increased, the two P-waves gradually separated. qS1 wave velocity was larger than qP4, which matches the results in Figure 8. The direct waveform was identical to the Ricker wavelet used as the excitation source. Because the thickness of the coal seam is less than one-quarter of the wavelength, the reflected waveform becomes a superposition of two Ricker wavelets due to the seismic tuning effect. As shown in Figure 10, the finite-difference method employed in this study accurately described the reflection and transmission of seismic waves in a two-phase layered medium.

To study the influence of the fracture inclination angle on the response characteristics of seismic reflection waves, the fracture inclination angle was set to 45°, and the offset distance was set to 50 m. Figure 11A and B shows the variation of seismic amplitude with azimuth in z and x directions, respectively. It is difficult to observe the effects of microscopic rock-physics parameters on the seismic record. Therefore, seismic attributes were used to quantitatively describe the seismic response characteristics of different microscopic rock physics parameters. The constructed stratigraphic model had a single configuration and no lateral variation. The variation in seismic response characteristics mainly comes from changes in amplitude. Therefore, seismic attributes derived from amplitude features were mainly used to quantitatively characterize the indicated seismic response features, and seismic records before and after 10 ms relative to the peak of the reflected wave were used to calculate the amplitude attributes. The changes of the selected seismic attributes with respect to the seismic azimuth are shown in Figure 11C, where the dashed line is the seismic attribute extracted from the amplitude in the x direction, and the solid line is the seismic attribute extracted from the amplitude in the z direction. As the azimuth increased from 0° to 180°, the seismic amplitude attributes in the z and x directions showed a sinusoidal pattern over half a cycle, with the amplitude of change in the z direction much larger than that in the x direction.

To study how the response characteristics of CBM reservoir-reflected waves are influenced by the fracture inclination angle, the azimuth was set to 0°, and the offset distance was 50 m. The seismic response characteristics for different azimuths were simulated using the finite-difference method. The amplitudes in the x and z directions are shown in Figure 12A and B. Figure 12C illustrates the relationship between seismic amplitude attributes and fracture inclination angle. As the fracture inclination angle increased from 0° to 90°, the seismic amplitude attributes in the z and x directions exhibited cosine and sine patterns, respectively, with the amplitude of change in the z direction larger. The fracture inclination angle was set to 45°, the offset distance to 50 m, and the azimuth to 0°. The responses of the amplitude and seismic amplitude attributes in the z direction and x direction with respect to the porosity of the fracture are shown in Figure 13A-C, respectively. With increasing fracture porosity, seismic

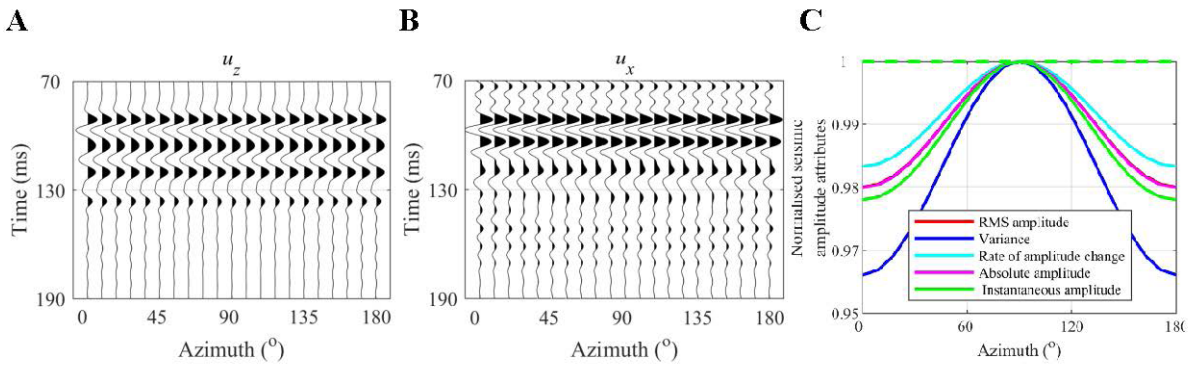



**Figure 11.** Response of seismic record and seismic amplitude attributes with respect to azimuth. (A) Displacement in the z-direction. (B) x-direction displacement. (C) Seismic amplitude attributes.

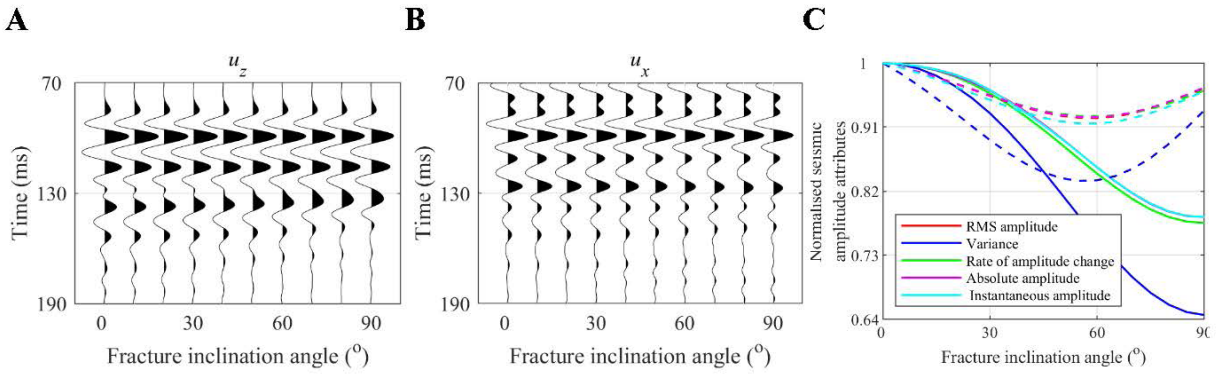



**Figure 12.** Response of seismic record and seismic amplitude attributes with respect to crack dip angle. (A) Displacement in the z-direction. (B) x-direction displacement. (C) Seismic amplitude attributes.

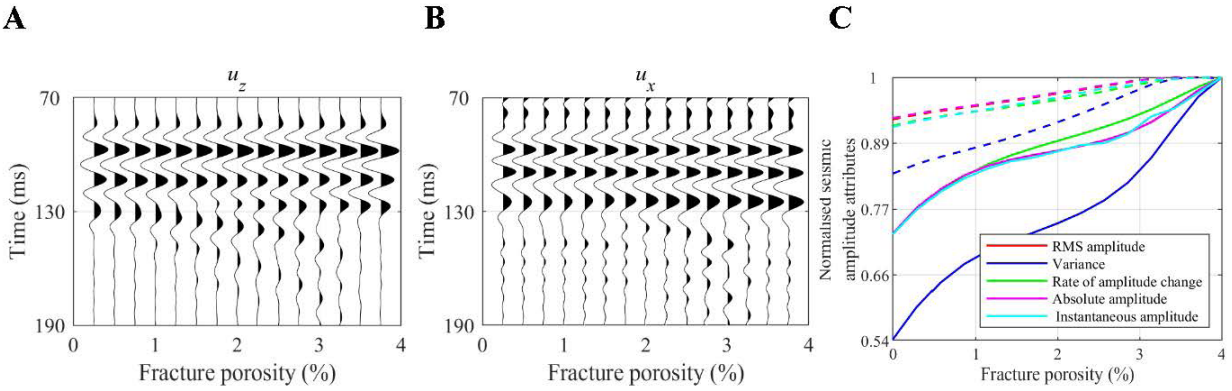



**Figure 13.** Response of seismic records and seismic amplitude attributes to crack porosity. (A) Displacement in the z-direction. (B) x-direction displacement. (C) Seismic amplitude attributes.

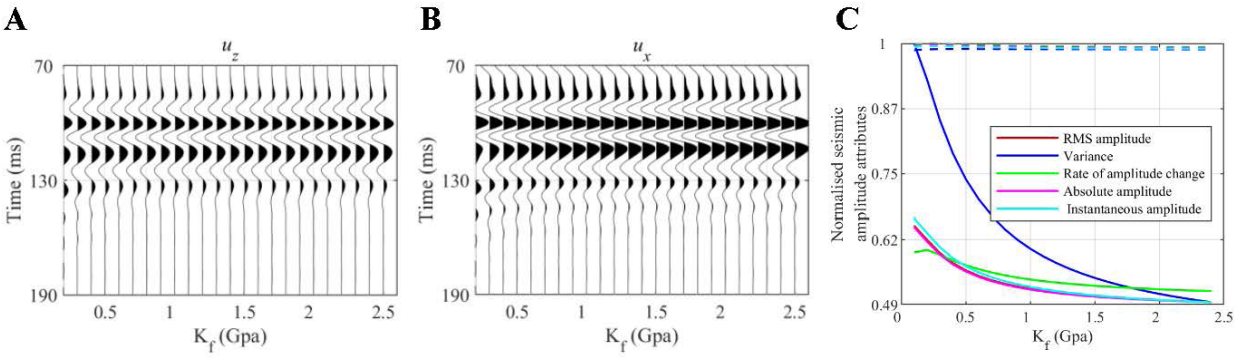



**Figure 14.** Response of seismic records and seismic amplitude attributes with respect to $K_f$. (A) Displacement in the z-direction. (B) x-direction displacement. (C) Seismic amplitude attributes.

amplitudes in both directions increased gradually but exhibited different patterns of change. The change in the z direction was larger than that x direction.

Figure 14 shows the effect of pore fluid equivalent bulk modulus on the seismic records of CBM reservoirs. As shown in Figure 14C, the seismic amplitude attributes in the x and z directions exhibit different trends as the bulk modulus of the mixed fluid increases gradually. The seismic amplitude attributes in the z direction gradually decreased, and the rate of decrease slowed; the change in the x direction is smaller than that in the z direction.

Combining Figures 11-14, it can be seen that with changes in azimuth, fracture parameters, and fluid parameters, the seismic records in the two-phase layered medium exhibited smaller changes, which were difficult to distinguish with the naked eye. The changes in seismic amplitude attributes were more obvious. x-direction seismic amplitude attributes were different from the z direction, and the changes of x-direction seismic amplitude attributes were smaller than those in the z direction. Comparing changes in seismic amplitude with the rock physics parameter in the z direction with those of RPP, it was found that the changes were the same. The RPP and seismic amplitude properties of a two-phase layered medium follow the same rules of change with rock physics parameters, which can reflect the medium's seismic AVO response characteristics to a certain extent. The correspondence between the derived RPP formula and seismic prediction results also demonstrates the formula's accuracy.

## 4. Discussion

This paper derives the RPP for a two-phase layered medium using Biot's theory. A rock physics model of CBM reservoirs was constructed based on their pore-structure characteristics, accounting for aligned fracture inclination angles and azimuths. Using this model, we predicted the response of RPP to changes in azimuth, aligned fracture inclination angle, aligned fracture porosity, pore fluid parameters, coalbed thickness, and the main frequency of the seismic wavelet in two-phase layered media. We simulated the propagation characteristics of seismic waves in CBM reservoirs using the finite-difference method with a rotated staggered grid. The results demonstrate the phenomena of fast and slow P-wave splitting and S-wave splitting in the solid and fluid phases of the medium. This indicates that the finite-difference method can demonstrate the propagation laws of seismic waves in bidirectionally anisotropic media. To investigate the seismic response characteristics of CBM reservoirs, we constructed a single-layer stratigraphic model and predicted seismic responses for different azimuths, fracture inclination angles, fracture porosity, and pore fluid bulk moduli. To highlight the rules governing changes in seismic features in relation to rock physics parameters, seismic amplitude attributes were used to quantify the seismic response. The seismic amplitude attributes in the z and x directions differed from the rock physics attributes, with the change in the z direction being larger than in the x direction. The seismic amplitude attributes in the z direction were similar to the RPP of the two-phase laminar medium. The RPP of the two-phase layered medium exhibited a high degree of similarity with the variation of seismic amplitude attributes. This study validated the seismic AVO response law in complex media and quantified the seismic response characteristics of CBM reservoirs under different seismic observation azimuths, fracture inclination angles, and other multi-scale rock physics parameters. It also supported accurate reservoir prediction and high-precision lithological inversion of CBM reservoirs.

## 5. Conclusion

This study investigated the coupled effects of multi-scale parameters—fracture dip angle, fluid, frequency, and bed thickness—on the azimuthal AVO response in CBM deposits. The theoretical reflection coefficient model elucidates the independent mechanisms of each parameter, whilst full-waveform simulation validates their robustness in actual complex wavefields. It further explores application pathways for extracting these patterns from data via seismic attributes. This methodology provides a physically well-founded solution for precise seismic interpretation of complex reservoirs.

It should be noted that due to the unique characteristics of CBM reservoirs, coal seam AVO responses are influenced by multiple factors. These influencing factors stem from the complex physical process of seismic wave incidence and reflection within coal seams. This study simulated the impact of these factors on reflected wave amplitudes. Given the complexity of these influencing factors, reservoir prediction based on forward modeling principles requires integrating appropriate nonlinear reservoir prediction methods. This issue should be addressed in future work.

## Acknowledgments

None.

## Funding

This study was supported by the National Natural Science Fund Projects (42274165,42474146), CNPC Innovation Fund (2024DQ02-0133), and the Fundamental Research Funds for the Central Universities (2024ZKPYDC03).

## Conflict of interest

Fei Gong and Qiang Guo are Editorial Board Members of this journal, but were not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* ZhaoJi Zhang, Fei Gong
*Formal analysis:* ZhaoJi Zhang
*Investigation:* ZhaoJi Zhang, Fei Gong
*Methodology:* ZhaoJi Zhang, Hao Li
*Visualization:* ZhaoJi Zhang, Hao Li
*Writing–original draft:* ZhaoJi Zhang
*Writing–review & editing:* GuanGui Zou, Qiang Guo, GuoWei Zhu

## Availability of data

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## References


1. Pan J, Ge T, Liu W, *et al*. Organic matter provenance and accumulation of transitional facies coal and mudstone in Yangquan, China: Insights from petrology and geochemistry. *J Nat Gas Sci Eng*. 2021;94:104076.

   doi: 10.1016/j.jngse.2021.104076

2. Zhang D, He H, Ren Y, *et al*. A mini review on biotransformation of coal to methane by enhancement of chemical pretreatment. *Fuel*. 2022;308:121961.

   doi: 10.1016/j.fuel.2021.121961

3. Guo Z, Li X. Azimuthal AVO signatures of fractured poroelastic sandstone layers. *Explor Geophys*. 2017;48(1):56-66.

   doi: 10.1071/eg15050

4. Hu H, Xia K, Hilterman F, Zhang Y. Amplitude-Versus-Angle Analysis of Local Angle-Domain Common Image Gathers With Prestack Gaussian-Beam Migration of Seismic Data. *IEEE Trans Geosci Remote Sens*. 2020;58(8):5969-5975.

   doi: 10.1109/TGRS.2020.2973654

5. Wandler A, Evans B, Link C. AVO as a fluid indicator: A physical modeling study. *Geophysics*. 2007;72(1):C9-C17.

   doi: 10.1190/1.2392817

6. Xu C, Chen C, Deng J, Di B, Wei J. A physical modelling study of the reservoir scale effect on amplitude variation with incidence angle. *J Geophys Eng*. 2019;16(6):1084-1093.

   doi: 10.1093/jge/gxz077

7. Xiao K, Zou C, Shang J, Yang Y, Zhang H. Numerical simulation of AVO response characteristics from pore-filling gas hydrate in Qilian mountain permafrost, China. *Arab J Geosci*. 2019;12(12):379.

   doi: 10.1007/s12517-019-4548-0

8. Gavin LJ, Lumley D. The effects of azimuthal anisotropy on 3D and 4D seismic amplitude variation with offset responses. *Geophysics*. 2019;84(6):C251-C267.

   doi: 10.1190/geo2018-0450.1

9. Liu H, Ding P, Li X. Physical modeling of seismic responses in thin interbedded reservoirs with horizontal fracture. *Chin J Geophys*. 2021;64(8):2927-2940.

   doi: 10.6038/cjg2021O0167

10. Xi Y, Yin X. Seismic Response Models and the AVO Simulation of Coal-Bearing Reservoirs. *Minerals*. 2022;12(7):836.

    doi: 10.3390/min12070836

11. Kumar D, Zhao Z, Foster DJ, Dralus D, Sen MK. Frequency-dependent AVO analysis using the scattering response of a layered reservoir. *Geophysics*. 2020;85(2):N1-N16.

    doi: 10.1190/geo2019-0167.1

12. Ahmed N, Weibull WW, Quintal B, Grana D, Bhakta T. Frequency-dependent AVO inversion applied to physically based models for seismic attenuation. *Geophys J Int*. 2022;233(1):234-252.

    doi: 10.1093/gji/ggac461

13. Ouyang F, Liu XZ, Wang B, *et al*. The applicability and underlying factors of frequency-dependent amplitude-versus-offset (AVO) inversion. *Pet Sci*. 2023;20(4):2075-2091.

    doi: 10.1016/j.petsci.2023.02.011

14. Tian Y, Stovas A, Gao J, Meng C, Yang C. Frequency-Dependent AVO Inversion and Application on Tight Sandstone Gas Reservoir Prediction Using Deep Neural Network. *IEEE Trans Geosci Remote Sens*. 2023;61:1-13.

    doi: 10.1109/tgrs.2023.3328183

15. Chen SQ, Li XY, Wang SX. The analysis of frequency-dependent characteristics for fluid detection: a physical model experiment. *Appl Geophys*. 2012;9(2):195-206.

    doi: 10.1007/s11770-012-0330-8

16. Si W, Di B, Wei J. Seismic response variation of tight gas sand for uniform and patchy saturation patterns. *J Appl Geophys*. 2015;116:167-172.

    doi: 10.1016/j.jappgeo.2015.03.010

17. Wang X, Pan D. Application of AVO attribute inversion technology to gas hydrate identification in the Shenhu Area, South China Sea. *Mar Pet Geol*. 2017;80:23-31.

    doi: 10.1016/j.marpetgeo.2016.11.015

18. Yadav A, Mondal S, Yadav J, Chakraborty S. Direct hydrocarbon indicator analysis to predict reservoir in Deepwater Krishna-Godavari basin: a case study. *Explor Geophys*. 2023;54(6):625-635.

doi: 10.1080/08123985.2023.2236119

19. Yang X, Cao S, Guo Q, Kang Y, Yu P, Hu W. Frequency-Dependent Amplitude Versus Offset Variations in Porous Rocks with Aligned Fractures. *Pure Appl Geophys*. 2016;174(3):1043-1059.

doi: 10.1007/s00024-016-1423-8

20. Castoro A, White RE, Thomas RD. Thin-bed AVO: Compensating for the effects of NMO on reflectivity sequences. *Geophysics*. 2001;66(6):1714-1720.

doi: 10.1190/1.1487113

21. Jin Z, Chapman M, Papageorgiou G, Wu X. Impact of frequency-dependent anisotropy on azimuthal P-wave reflections. *J Geophys Eng*. 2018;15(6):2530-2544.

doi: 10.1088/1742-2140/aad882

22. Saeed W, Zhang H, Guo Q, *et al*. An integrated petrophysical-based wedge modeling and thin bed AVO analysis for improved reservoir characterization of Zhujiang Formation, Huizhou sub-basin, China: A case study. *Open Geosci*. 2020;12(1):256-274.

doi: 10.1515/geo-2020-0011

23. Shi J, Wang S, Zhang H, *et al*. A novel method for formation evaluation of undersaturated coalbed methane reservoirs using dewatering data. *Fuel*. 2018;229:44-52.

doi: 10.1016/j.fuel.2018.04.144

24. Wang P, Yao Z, Peng S, Hu S. A novel approximate formula on AVAZ of reflection and transmission qP in TTI media. *J Pet Sci Eng*. 2020;192:107280.

doi: 10.1016/j.petrol.2020.107280

25. Xue J, Gu H, Cai C. Model-based amplitude versus offset and azimuth inversion for estimating fracture parameters and fluid content. *Geophysics*. 2017;82(2):M1-M17.

doi: 10.1190/geo2016-0196.1

26. Shi B, Wang P. Research on Stability Control of Shields at Working Face with Large Dip Angle. *Energies*. 2023;16(15):5813.

doi: 10.3390/en16155813

27. Tu H, Tu S, Yuan Y, Wang F, Bai Q. Present situation of fully mechanized mining technology for steeply inclined coal seams in China. *Arab J Geosci*. 2014;8(7):4485-4494.

doi: 10.1007/s12517-014-1546-0

28. Elkibbi M, Yang M, Rial JA. Crack-induced anisotropy models in The Geysers geothermal field. *Geophys J Int*. 2005;162(3):1036-1048.

doi: 10.1111/j.1365-246x.2005.02697.x

29. Kolesnikov Y, Fedin KV, Ngomayezwe L. Compression waves reflection from the low-velocity azimuthally anisotropic medium: a physical model study. *Geophys J Int*. 2020;221(2):1320-1326.

doi: 10.1093/gji/ggaa031

30. Zhang J, Fan X, Huang Z, Liu Z, Fan Z, Liu L. In situ stress determination in isotropic and anisotropic rocks and its application to a naturally fractured reservoir. *Geomech Geophys Geo-Energy Geo-Resour*. 2023;9(1):80.

doi: 10.1007/s40948-023-00584-6

31. Huang Y, Wei M, Malekian R, Zhen X. CBM Reservoir Rock Physics Model and Its Response Characteristic Study. *IEEE Access*. 2017;5:5837-5843.

doi: 10.1109/access.2017.2687882

32. Wu H, Guo J, Ji G, Huang Y, Ding H, Lin P. Estimating the anisotropy of the vertical transverse isotropy coal seam by rock physics model–based inversion. *Geophys Prospect*. 2024;72(5):2064-2075.

doi: 10.1111/1365-2478.13470

33. Gong F, Zou G, Zhang Z, Peng S, Wang G, Chen H. An anisotropic rock physics modeling for the coalbed methane reservoirs and its applications in anisotropy parameter prediction. *J Appl Geophys*. 2024;225:105381.

doi: 10.1016/j.jappgeo.2024.105381

34. Mavko G, Mukerji T, Dvorkin J. *The Rock Physics Handbook*. 2nd ed. Cambridge University Press; 2010.

doi: 10.1017/cbo9780511626753

35. Zhao M, Jin Y, Liu X, Zheng J, Liu S. Characterizing the Complexity Assembly of Pore Structure in a Coal Matrix: Principle, Methodology, and Modeling Application. *J Geophys Res Solid Earth*. 2020;125(12):e2020JB020110.

doi: 10.1029/2020jb020110

36. Gong F, Huang A, Kang W, *et al*. The influence of lamination and fracture on the velocity anisotropy of tectonic coals. *Geophysics*. 2024;89(6):MR355-MR365.

doi: 10.1190/geo2024-0033.1

37. Alkhimenkov Y, Quintal B. A simple and accurate model for attenuation and dispersion caused by squirt flow in isotropic porous rocks. *Geophysics*. 2023;89(1):MR1-MR10.

doi: 10.1190/geo2023-0049.1

38. Mavko G, Nolen-Hoeksema RC. Estimating seismic velocities at ultrasonic frequencies in partially saturated rocks. *Geophysics*. 1994;59(2):252-258.

doi: 10.1190/1.1443587

39. Biot MA. Theory of Propagation of Elastic Waves in a Fluid-Saturated Porous Solid. I. Low-Frequency Range. *J Acoust Soc Am*. 1956;28(2):168-178.

doi: 10.1121/1.1908239

40. Carcione JM, Kosloff D, Behle A. Long-wave anisotropy in stratified media; a numerical test. *Geophysics*. 1991;56(2):245-254.

doi: 10.1190/1.1443037

41. Chen T, Liu Y. Multi-component AVO response of thin beds based on reflectance spectrum theory. *Appl Geophys*. 2006;3(1):27-36.

doi: 10.1007/s11770-006-0004-5

42. Kennett BLN, Kerry NJ. Seismic waves in a stratified half space. *Geophys J Int*. 1979;57(3):557-583.

doi: 10.1111/j.1365-246x.1979.tb06779.x

43. Sidler R, Holliger K. Seismic reflectivity of the sediment-covered seafloor: effects of velocity gradients and fine-scale layering. *Geophys J Int*. 2010;181(1):521-531.

doi: 10.1111/j.1365-246x.2010.04519.x

44. Wang Y, Yang C, Lu J. Dilemma faced by elastic wave inversion in thinly layered media. *Chin J Geophys*. 2018;61(3):1118-1135.

doi: 10.6038/cjg2018L0404

45. Liu Y, Schmitt DR. Amplitude and AVO responses of a single thin bed. *Geophysics*. 2003;68(4):1161-1168.

doi: 10.1190/1.1598108

46. Meissner R, Meixner E. Deformation of seismic wavelets by thin layers and layered boundaries. *Geophys Prospect*. 1969;17(1):1-27.

doi: 10.1111/j.1365-2478.1969.tb02069.x

47. Pan W, Innanen KA. AVO/AVF analysis of thin beds in elastic media. In: *SEG Technical Program Expanded Abstracts 2013*. Society of Exploration Geophysicists; 2013.

doi: 10.1190/segam2013-0587.1

48. Dai N, Vafidis A, Kanasewich ER. Wave propagation in heterogeneous, porous media; a velocity-stress, finite-difference method. *Geophysics*. 1995;60(2):327-340.

doi: 10.1190/1.1443769

49. Geertsma J, Smit DC. Some aspects of elastic wave propagation in fluid-saturated porous solids. *Geophysics*. 1961;26(2):169-181.

doi: 10.1190/1.1438855

50. Deresiewicz H, Skalak R. On uniqueness in dynamic poroelasticity. *Bull Seismol Soc Am*. 1963;53(4):783-788.

doi: 10.1785/bssa0530040783

51. Rokhlin SI, Wang YJ. Equivalent boundary conditions for thin orthotropic layer between two solids: Reflection, refraction, and interface waves. *J Acoust Soc Am*. 1992;91(4):1875-1887.

doi: 10.1121/1.403717

52. Tooley RD, Spencer TW, Sagoci HF. Reflection and transmission of plane compressional waves. *Geophysics*. 1965;30(4):552-570.

doi: 10.1190/1.1439622

53. Ursin B, Stovas A. Reflection and transmission responses of a layered isotropic viscoelastic medium. *Geophysics*. 2002;67(1):307-323.

doi: 10.1190/1.1451803

54. Mukerji T, Mavko G. Pore fluid effects on seismic velocity in anisotropic rocks. *Geophysics*. 1994;59(2):233-244.

doi: 10.1190/1.1443585

55. Zou G, Zeng H, Peng S, Zhou X, Satibekova S. Bulk density and bulk modulus of adsorbed coalbed methane. *Geophysics*. 2019;84(2):K11-K21.

doi: 10.1190/geo2018-0081.1

56. Saenger EH, Gold N, Shapiro SA. Modeling the propagation of elastic waves using a modified finite-difference grid. *Wave Motion*. 2000;31(1):77-92.

doi: 10.1016/s0165-2125(99)00023-2

57. Jianxiong L, Chunjiao M. An efficient FDTD implementation of the CFS-PML based on the ADE method and its validation along with the PLRC method in dispersive media. In: Hong W, Yang G, eds. *2008 International Conference on Microwave and Millimeter Wave Technology*. IEEE; 2008:766-769.

doi: 10.1109/icmmt.2008.4540510

58. Komatitsch D, Martin R. An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*. 2007;72(5):SM155-SM167.

doi: 10.1190/1.2757586

59. Martin R, Komatitsch D. An unsplit convolutional perfectly matched layer technique improved at grazing incidence for the viscoelastic wave equation. *Geophys J Int*. 2009;179(1):333-344.

doi: 10.1111/j.1365-246x.2009.04278.x

## Appendix

The boundary continuity conditions expressed in **Equations 12-17** may be represented in matrix form as follows:

$$\begin{bmatrix} u_x \\ u_z \\ M \\ \sigma_{zz}+s \\ \tau \\ P \end{bmatrix} = \begin{bmatrix} \frac{\partial \varphi_s}{\partial x} - \frac{\partial \psi_s}{\partial z} \\ \frac{\partial \varphi_s}{\partial z} + \frac{\partial \psi_s}{\partial x} \\ \rho_f (\frac{\partial \varphi_f}{\partial z} + \frac{\partial \psi_f}{\partial x}) \\ (C_{(3,:)} + C_{(7,:)}) \cdot \Im \\ C_{(5,:)} \cdot \Im \\ -\frac{1}{\phi} C_{(7,:)} \cdot \Im \end{bmatrix} \tag{A1}$$

where $C$ is the elastic stiffness matrix in a two-phase medium and $\Im$ is a strain-dependent quantity. The wave function expression is taken from **Equation 11** into **Equation 18**:

$$\begin{cases} \begin{bmatrix} u_x \\ u_z \\ m \end{bmatrix} = \begin{bmatrix} ikx_1 & ikx_2 & -ikz_3 & ikx_4 & ikx_5 & -ikz_6 \\ ikz_1 & ikz_2 & ikx_3 & ikz_4 & ikz_5 & ikx_6 \\ ib_1\rho_f kz_1 & ib_2\rho_f kz_2 & ib_3\rho_f kx_3 & ib_4\rho_f kz_4 & ib_5\rho_f kz_5 & ib_6\rho_f kx_6 \end{bmatrix} \cdot e \\ \begin{bmatrix} \sigma_{zz}+s \\ \tau \\ -P \end{bmatrix} = \begin{bmatrix} C_{(3,:)} + C_{(7,:)} \\ C_{(5,:)} \\ C_{(7,:)} / \phi \end{bmatrix} \cdot \Im \end{cases} \tag{A2}$$

where

$$\Im = \begin{bmatrix} -kx_1^2 & -kx_2^2 & kx_3kz_3 & -kx_4^2 & -kx_5^2 & kx_6kz_6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -kz_1^2 & -kz_2^2 & -kx_3kz_3 & -kz_4^2 & -kz_5^2 & -kx_6kz_6 \\ 0 & 0 & kx_3kz_3 - kz_3^2 & 0 & 0 & kx_6kz_6 - kz_6^2 \\ -2kx_1kz_1 & -2kx_2kz_2 & kz_3^2 - kx_3^2 & -2kx_4kz_4 & -2kx_5kz_5 & kz_6^2 - kx_6^2 \\ 0 & 0 & kx_3^2 - kx_3kz_3 & 0 & 0 & kx_6^2 - kx_6kz_6 \\ -b_1kx_1^2 - b_1kz_1^2 & -b_2kx_2^2 - b_2kz_2^2 & 0 & -b_4kx_4^2 - b_4kz_4^2 & -b_5kx_5^2 - b_5kz_5^2 & 0 \end{bmatrix} e \tag{A3}$$

$$e = \begin{bmatrix} \mathrm{e}^{izk_{z1}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathrm{e}^{izk_{z2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathrm{e}^{izk_{z3}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{e}^{izk_{z4}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathrm{e}^{izk_{z5}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathrm{e}^{izk_{z6}} \end{bmatrix} \cdot \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \end{bmatrix} \mathrm{e}^{i(\delta x - tw)} \tag{A4}$$

The relationship between the continuity condition and the solid phase amplitude in the n-th layer can be simplified as:

$$\begin{bmatrix} u_x^{(n)} \\ u_z^{(n)} \\ M^{(n)} \\ \sigma_{zz}^{(n)} + s^{(n)} \\ \tau_{xz}^{(n)} \\ -P^{(n)} \end{bmatrix} = H_n \cdot \begin{bmatrix} A_1^{(n)} \\ A_2^{(n)} \\ A_3^{(n)} \\ A_4^{(n)} \\ A_5^{(n)} \\ A_6^{(n)} \end{bmatrix} \tag{A5}$$

where $H_n$ is the relationship between the amplitude and the boundary continuity condition in the n-th layer of the medium, which is determined by Equations A1-A5. A recursive formula was developed to determine the quantitative relationship of the continuity condition at the top and bottom boundaries of the laminated medium:

$$\begin{bmatrix} u_x^{(n)} \\ u_z^{(n)} \\ M^{(n)} \\ \sigma_{zz}^{(n)} + s^{(n)} \\ \tau_{xz}^{(n)} \\ -P^{(n)} \end{bmatrix} = [H_{m-1}]_{z=Z(m-1)} \cdot d_{m-1} \cdots d_3 \cdot [H_2]_{z=Z(1)}^{-1} \cdot \begin{bmatrix} u_x^{(1)} \\ u_z^{(1)} \\ M^{(1)} \\ \sigma_{zz}^{(1)} + s^{(1)} \\ \tau_{xz}^{(1)} \\ -P^{(1)} \end{bmatrix} \tag{A6}$$

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Mitigating numerical dispersion in full-waveform inversion imaging: A preconditioned optimization approach for finite-difference weights

**Ganglin Lei[1,2,3,4,5], Jianping Huang[6]*, Wensheng Duan[1,2,3,4,5], Chao Chen[1,2,3,4,5], Zhenwen Liu[1,2,3,4,5], Chang Zhou[1,2,3,4,5], and Weiting Peng[6]**

[1]R&D Center for Ultra-Deep Complex Reservoir Exploration and Development, China National Petroleum Corporation, Korla, Xinjiang, China

[2]Xinjiang Engineering Research Center for Ultra-deep Complex Reservoir Exploration and Development, China National Petroleum Corporation, Korla, Xinjiang, China

[3]Xinjiang Key Laboratory of Ultra-deep Oil and Gas, China National Petroleum Corporation, Korla, Xinjiang, China

[4]China National Energy Administration, State Energy Key Laboratory of Carbonate Oil and Gas, Korla, Xinjiang, China

[5]Key Laboratory of Carbonate, China National Petroleum Corporation, Korla, Xinjiang, China

[6]School of Geosciences, China University of Petroleum (East China), Qingdao, Shandong, China

## Abstract

Full-waveform inversion (FWI) imaging is a high-resolution seismic imaging technique that directly produces subsurface images by inverting the full recorded wavefield. However, its reliability is often limited by numerical dispersion errors arising from finite-difference (FD) forward modeling. One key approach for reducing dispersion is to optimize the FD coefficients using an optimization algorithm. However, conventional methods for optimizing FD weights focus only on reducing spatial dispersion, which can weaken numerical stability, especially when using large time steps (i.e., high Courant–Friedrichs–Lewy [CFL] numbers). To address this issue, we introduce a new optimization approach that improves both simulation accuracy and stability. The proposed method combines error functions from both the time–space domain and the spatial domain into a single adaptive objective function. A dynamic weighting factor, which depends on the CFL number, facilitates a trade-off between accuracy and stability of the optimal FD weights. We also use the seismic wavelet spectrum as prior information to constrain the optimization. The optimization problem is solved by the least-squares method. In the theoretical test, the proposed weights significantly reduce wavefield simulation errors across a wide range of wavenumbers, with a higher CFL number than conventional approaches. When applied to FWI, this method reduces phase distortion and local minima in the objective function. In a test using the Marmousi model at 40 Hz, our approach produced clear and continuous deep structures, closely matching results from dispersion-free benchmarks. In contrast, conventional methods failed due to severe dispersion. This work provides a more robust numerical foundation for high-frequency FWI imaging by improving both accuracy and stability.

***Keywords:*** Finite-difference weights; Full-waveform inversion; Full-waveform inversion imaging; Dispersion error

***Corresponding author:**
Jianping Huang
(jphuang@upc.edu.cn)

**Citation:** Lei G, Huang J, Duan W, *et al.* Mitigating numerical dispersion in full-waveform inversion imaging: A preconditioned optimization approach for finite-difference weights. *J Seismic Explor.* 2026;35(2):025440096. doi: 10.36922/JSE025440096

**Received:** October 29, 2025

**Revised:** December 21, 2025

**Accepted:** December 22, 2025

**Published online:** March 6, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Full-waveform inversion (FWI), initially proposed by Lailly[1] and Tarantola,[2] and extensively reviewed by Virieux and Operto,[3] is an elegant and natural way to handle full-wavefield data with a proper modeling engine. The power of FWI not only enhances velocity model building but also opens new avenues for seismic imaging.[4-6] By leveraging the full wavefield and high-resolution physical property models recovered by FWI, it becomes possible to derive reflectivity and structural information directly from these models, offering an alternative to conventional migration techniques.[7,8]

FWI imaging has advanced rapidly in recent years. Lu[9] demonstrated that the one-dimensional derivative of a high-resolution FWI velocity model yields a reflectivity estimate that delivers enhanced imaging compared to conventional techniques. To better image dipping events, Huang *et al.*[10] applied a three-dimensional derivative aligned with the normal to the dominant dip direction. In a different approach, Operto *et al.*[11] performed a direct reflectivity inversion within a multiparameter FWI framework.

However, the highly non-linear nature of the FWI objective function makes it highly susceptible to local minima, with "cycle-skipping" being a central challenge to its convergence. To mitigate this issue, various strategies have been developed, primarily focusing on expanding the inversion search space and enhancing the convexity of the objective function. For instance, reflection FWI circumvents cycle-skipping by separating tomographic and migration modes, preferentially using reflected waves to construct long-wavelength background velocity models.[12] Furthermore, physics-based low-frequency extrapolation or the use of deep learning networks to directly build initial models have proven effective in enlarging the convergence basin of FWI.[13] Alternatively, extending the inversion search space beyond the single-scattering Born approximation provides a rigorous mathematical framework to mitigate the non-linearity issue.[14] On the other hand, designing more convex objective functions, such as envelopes or the Wasserstein distance, can mathematically reduce the risk of converging to local minima. In recent years, deep learning-based velocity inversion methods have shown great potential by learning complex non-linear mappings from data to models, bypassing the local minima issues inherent in conventional FWI.[13] Concurrently, stable numerical simulation forms the cornerstone of implementing these advanced algorithms. In this context, absorbing boundary conditions compatible with the finite-difference (FD) method, such as the residual perfectly matched layer, are crucial for suppressing boundary reflections and ensuring a clean wavefield.[15] Nevertheless, a fundamental and often overlooked problem remains: if the forward modeling engine itself contains significant errors due to numerical dispersion, it introduces non-physical, spurious local minima into the inversion process, which can undermine or even negate the effectiveness of the aforementioned advanced inversion techniques.

The accuracy of the numerical forward modeling in the inversion process is crucial for the reliability of the FWI results. Among various numerical methods, the FD scheme has become the predominant choice for large-scale FWI applications due to its computational efficiency and straightforward implementation.[16,17] FD simulations are often affected by numerical dispersion, which can significantly reduce the quality of the synthetic wavefield. The inherent errors of the FD scheme arise from the truncation of higher-order terms in the FD approximation, leading to spatial discretization errors.[18,19] This dispersion phenomenon becomes increasingly severe at higher frequencies. Unfortunately, reflectivity imaging—essentially a high-frequency inversion process—requires the numerical simulation of high-frequency seismic wavefields during FWI imaging. Such a high-frequency wavefield simulation tends to exacerbate numerical dispersion.

To address these problems, a common strategy to mitigate this dispersion error is the optimization of FD weights with broad bandwidth coverage.[20-22] Various optimization approaches have been developed to compute FD weights. One approach involves optimization under different norms, including the $L_\infty$-norm,[20-29] the $L_2$-norm,[30-36] and the $L_1$-norm.[37] Another strategy uses window functions to truncate the wavefield's pseudospectrum for coefficient derivation.[22,38-42]

An ideal FD scheme would be entirely free from numerical dispersion across all frequencies, perfectly preserving the dispersion relation. Although this ideal is not yet attainable, two practical criteria are widely considered: maximizing the bandwidth within a given error tolerance, or minimizing the error within a given bandwidth.[28] Due to its equiripple characteristic, the optimized FD weights delivered by the Remez exchange algorithm can maximize the bandwidth under a given error tolerance.[27,28,43,44] In contrast, due to its $L_1$-norm objective function, the alternating direction method of multipliers approach is effective at minimizing the numerical error within a given bandwidth.[37]

However, since conventional optimization of FD weights does not account for the influence of simulation parameters, the resulting "optimized" weights may, in some cases, introduce even greater simulation inaccuracies. This

phenomenon indicates a fundamental limitation: existing methods minimize the dispersion relation error rather than the actual wavefield simulation error. Although Peng and Huang[45] incorporated wavefield information to improve FD simulations by reducing the spatial dispersion relation error, their approach overlooked numerical stability considerations. Optimization schemes based solely on the spatial dispersion relation often reduce the stability of the FD simulation. Moreover, in their computational process, they used a double-loop mechanism combined with an exhaustive search method to compute the optimized weights, which significantly increased computational resource consumption.

Furthermore, practical applications frequently face challenges related to stability constraints, especially when spatial or temporal sampling rates exceed the stability condition. Most of these optimization approaches focus solely on minimizing the error of the spatial dispersion relation. This way of optimizing the FD weights often compromises the stability of the FD method and fails to meet the demands of practical applications. In a previous study, Liu and Sen[46] introduced a method that uses the error in the spatial–temporal dispersion relation as the objective function to optimize FD weights, which was shown to extend the stability limit of the FD scheme.

In this study, we propose a novel optimization method for FD weights that combines the dispersion relation in the space domain and the time–space domain. By incorporating wavefield simulation parameters as prior information, we formulate an $L_2$-norm constrained optimization problem and obtain a global solution via the least-squares method. Through theoretical analysis and numerical tests, we demonstrate the advantages of the proposed method in reducing dispersion error and enhancing stability. Furthermore, we investigate the impact of FD dispersion errors on the FWI objective function and subsequently evaluate the FWI imaging performance using both the proposed and conventional methods on simple and complex models. Our results demonstrate that the new method effectively suppresses numerical dispersion in FWI, mitigates cycle-skipping issues, and delivers high-resolution images.

## 2. Theory of optimized FD weights

The two-dimensional acoustic wave equation is expressed as **Equation 1**:

$$\frac{1}{v^2}\frac{\partial^2 u_0^0}{\partial t^2}=\frac{\partial^2 u_0^0}{\partial x^2} \tag{1}$$

where $u_q^p$ denotes the scalar pressure wavefield, with the notation shown in **Equation 2**:

$$u_q^p = u\mathrm{e}^{ik\left(x+qh\right)-i\omega\left(t+p\tau\right)} \tag{2}$$

where $e$ is the base of the natural logarithm, $u$ denotes a constant amplitude of the wavefield, $k$ is the wavenumber, and $\omega$ is the angular frequency. Here, $h$ represents the spatial grid size, $t$ is the time step, and $v$ stands for the velocity.

A conventional explicit second-order FD approximation for the temporal derivative is given by **Equation 3**:

$$\frac{\partial^2 u_0^0}{\partial t^2}\approx\frac{1}{\tau^2}\left[-2u_0^0+u_0^1+u_0^{-1}\right] \tag{3}$$

On a uniform grid, a 2M[th]-order FD discretization of the spatial derivative can be written as **Equation 4:**[29,38]

$$\frac{\partial^2 u_0^0}{\partial x^2}\approx\frac{1}{h^2}\left\{c_0u_0^0+\sum_{m=1}^{M}c_m\left[u_m^0+u_{-m}^0\right]\right\} \tag{4}$$

where $c_0$ and $c_m$ are FD weights, and 2M denotes the order of the FD scheme. Substituting **Equations 3** and **4** into **1** and rearranging it, we can obtain **Equation 5**:

$$\frac{1}{\tau^2}\left[-2u_0^0+u_0^1+u_0^{-1}\right]\approx\frac{v^2}{h^2}\left\{c_0u_0^0+\sum_{m=1}^{M}c_m\left[u_m^0+u_{-m}^0\right]\right\} \tag{5}$$

Substituting **Equation 2** into the above expression transforms it into the frequency–wavenumber domain, as shown in **Equation 6**:

$$\frac{1}{\tau^2}\left[-2+2\cos\left(\omega\tau\right)\right]U\left(k,\omega\right)\approx\frac{v^2U\left(k,\omega\right)}{h^2}\left[c_0+2\sum_{m=1}^{M}c_m\cos\left(mkh\right)\right] \tag{6}$$

where $U(k)$ denotes the spectrum of the seismic wavefield $u(x)$. When $k$=0, the FD weights satisfy $c_0=-2\sum_{m=1}^{M}c_m$. Using this relation and the identity $v=\frac{\omega}{k}$, **Equation 6** can be reformulated as **Equation 7**:

$$r^{-2}\left[-2+2\cos\left(rkh\right)\right]U\left(k\right)\approx U\left(k\right)2\sum_{m=1}^{M}c_m\left[\cos\left(mkh\right)-1\right] \tag{7}$$

where $r=\frac{v\tau}{h}$, $r$ denotes the Courant–Friedrichs–Lewy (CFL) number.

The value 0.505 represents the maximum stable CFL number ($r_{\max}$) for the conventional FD weights when M=6, which serves as a key parameter for evaluating numerical stability. Furthermore, as the order of the FD scheme increases (increasing M), the $r_{\max}$ continues to decrease.

The FD method introduces numerical errors in wave equation simulation due to the truncation of the Taylor series in its approximation. The specific form of this error is given by its expression in the frequency–wavenumber domain, as shown in **Equation 8**:

$$\varepsilon = U(k)\left\{\frac{2}{rkh}\arcsin\sqrt{r^2\sum_{m=1}^{M}c_m\sin^2\left(mkh/2\right)}-1\right\} \tag{8}$$

where $\frac{2}{rkh}\sin^{-1}\sqrt{r^2\sum_{m=1}^{M}c_m\sin^2\left(mkh/2\right)}-1$ represents the phase velocity error formula. If we directly construct the objective function based on the phase velocity error, the non-linear nature of the inverse trigonometric function would make it difficult to solve for the optimized FD weights. To circumvent this, we derived a simplified objective function from **Equation 7** that avoids the arcsin function non-linearity, as shown in **Equation 9**:

$$\Omega = U(k)\left\{\frac{r^2\sum_{m=1}^{M}c_m\left[\cos\left(mkh\right)-1\right]}{\left[\cos\left(rkh\right)-1\right]}-1\right\} \tag{9}$$

Thus, the numerical simulation error arising from the discretization can be expressed as **Equation 10**:

$$E = \int_0^{k_{\max}} U(k)\Phi(k)dk \tag{10}$$

where $E$ denotes the cumulative simulation error in the time-space domain, and $\Phi(k)$ represents the dispersion relation in the time–space domain, defined as $\frac{r^2\sum_{m=1}^{M}c_m\left[\cos\left(mkh\right)-1\right]}{\left[\cos\left(rkh\right)-1\right]}-1$. Here, is the maximum wavenumber. According to the Nyquist theorem, the maximum value of $kh$ is . Therefore, **Equation 10** can be rewritten as **Equation 11**:

$$E = \int_0^{\frac{\pi}{h}} U(k)\Phi(k)dk \tag{11}$$

Optimizing the FD weights based on this time–domain functional error can significantly improve the stability of the FD scheme. However, such optimization alone is insufficient to fully suppress spatial dispersion errors. To address this limitation, it is necessary to introduce a regularization term into the objective function that specifically constrains spatial dispersion. In this study, we incorporated the relative error of the spatial dispersion relation as an additional constraint to effectively reduce spatial dispersion during numerical simulation.

The spatial dispersion relation is derived from the spatial second derivative in the wave equation and its FD approximation, mathematically expressed as **Equation 12**:

$$\frac{\partial^2 u_0^0}{\partial x^2} \approx \frac{1}{h^2}\left\{c_0u_0^0+\sum_{m=1}^{M}c_m\left[u_m^0+u_{-m}^0\right]\right\} \tag{12}$$

By substituting the plane wave solution, it can be simplified into the following form shown in **Equation 13**:

$$\left(kh\right)^2U(k) \approx 2\sum_{m=1}^{M}c_m\left[\cos\left(mkh\right)-1\right]U(k) \tag{13}$$

Minimizing the spatial dispersion error ensures that the FD scheme closely approximates the true, continuous spatial second derivative, thereby yielding optimized FD weights that effectively reduce spatial dispersion.

The dispersion relation in the time–space domain, on the other hand, involves approximating the FD scheme of the temporal second derivative using the FD scheme of the spatial second derivative. Minimizing the dispersion error in the time–space domain aims to align these two schemes more closely. In this case, the objective of optimizing FD weights does not accurately approximate the true continuous spatial second derivative by minimizing the dispersion error in the time–space domain. Consequently, optimizing the temporal dispersion error alone does not contribute to the reduction of spatial dispersion.

To simultaneously extend the stability limit and minimize spatial dispersion, we proposed a novel adaptive weighting method for constructing the objective function. The core idea is to dynamically adjust, based on the current CFL number $r$ relative to the stability limit $r_{\max}$, the weighting in the objective function between the time–space domain dispersion error and the spatial domain dispersion error. This enables the optimization process to ensure numerical stability at high CFL numbers while maintaining spatial simulation accuracy at low CFL numbers. The new objective function is defined as follows in **Equation 14**:

$$\begin{aligned} J &= \int_0^{\frac{\pi}{h}}\left\|U(k)\Phi(k)\right\|_2^2dx+\lambda(r)\int_0^{\frac{\pi}{h}}\left\|U(k)\phi(k)\right\|_2^2dk \\ &= A+\lambda(r)B \end{aligned} \tag{14}$$

where $\left\|\bullet\right\|_2^2$ represents the $L_2$-norm function, and $J$ represents the total objective function value. $A$ denotes the time–space domain dispersion, which directly influences the stability of the numerical simulation. Minimizing $A$ is crucial for ensuring that the time-stepping algorithm remains stable at larger CFL numbers. In **Equation 14**,

$\phi(k)=\frac{2\sum_{m=1}^{M}c_m\left[\cos(mkh)-1\right]}{(kh)^2}$, and $B$ represents the spatial domain dispersion, which controls the accuracy of the simulation.

Minimizing $B$ effectively suppresses numerical dispersion and ensures the accuracy of wavefield modeling. $\lambda(r)$ is the adaptive weighting function. It is the core of this method. This function depends on the CFL number $r$ and dynamically balances the weights of $A$ and $B$ during the optimization process, as shown in **Equation 15**:

$$\lambda(r)=\lambda_0\left[1-\frac{1}{1+e^{-p(r-r_{tr})}}\right] \quad (15)$$

where $\lambda(r)$ is the adaptive weighting function. The weighting function $\lambda(r)$ is designed to prioritize numerical accuracy at low CFL numbers while ensuring numerical stability at high CFL numbers. We employed a sigmoid function for$\lambda(r)$, and the design rationale for its three core hyperparameters is detailed below:

(i) Base weight ($\lambda_0$ = 1,000): This parameter is assigned a very large value to ensure that at low CFL numbers ($r < r_{tr}$), the objective function is dominated by the spatial dispersion error $B$. This establishes an "accuracy-priority" optimization strategy, guaranteeing that the optimized FD weights suppress spatial numerical dispersion to the greatest extent possible under conditions where stability is not a constraint, thereby preserving wavefield fidelity.

(ii) Transition Center ($r_{tr}$): For example, when using the optimized FD weights (M=6) obtained by a conventional optimization algorithm, such as the least-square method, the stability limit $r_{max} < 0.505$, and $r_{tr}$=0.5050. It acts as a proactive buffer threshold, marking the critical point where the optimization strategy shifts from "accuracy-priority" to "stability-priority." By triggering this transition preemptively, the algorithm begins to adapt before nearing the theoretical stability boundary, thereby enhancing the robustness of the method.

(iii) Steepness factor ($p$ = 100): A relatively large steepness factor ensures that the weighting function $\lambda(r)$ changes rapidly and sharply around the transition center $r_{tr}$. This "switch-like" behavior forces the optimizer's focus to shift decisively from minimizing dispersion error in the space domain to minimizing that in the time–space domain once the CFL number enters the critical region, enabling an efficient search for coefficients that satisfy the stringent stability requirements at high CFL numbers.

These three hyperparameters work in concert to form an intelligent optimization framework. They empower our method to automatically and smoothly make an optimal trade-off between accuracy and stability in response to the CFL number, which is the key to achieving higher computational efficiency without sacrificing numerical accuracy.

In **Equation 14**, the functional incorporates the seismic wavefield spectrum. Directly obtaining this spectrum generally necessitates numerical forward modeling—a process often computationally prohibitive for practical applications. Given that the wavenumber spectrum of the wavefield in forward modeling is largely governed by the frequency spectrum of the seismic wavelet, we approximated the preconditioning operator using the wavelet spectrum.

Using a wavelet to approximate the wavefield function constitutes an incomplete representation, which may introduce errors under certain conditions, such as in complex media. We conducted forward modeling using 20 Hz and 40 Hz seismic wavelets in a complex medium, and compared the amplitude spectra of the seismic data with those of the source wavelet. As illustrated in Figures 1 and 2, under both 20 Hz and 40 Hz conditions, discernible deviations were observed between the amplitude spectrum of the seismic data (blue curve) and that of the seismic wavelet (red curve) across both high- and low-frequency ranges. This indicates that the wavelet spectrum does not fully characterize the frequency response of the seismic wavefield. Nevertheless, such an approximation error remains well within acceptable limits for our algorithm. Therefore, we adopted the wavelet

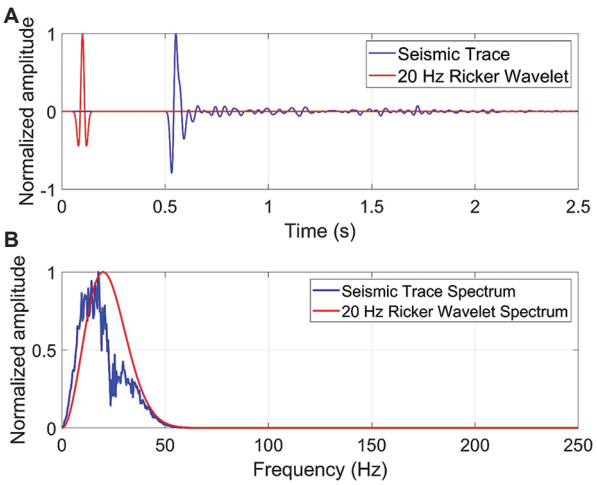



**Figure 1.** Comparative analysis of the 20 Hz seismic data and theoretical Ricker wavelet in time and frequency domains. (A) Normalized time–domain comparison. (B) Normalized spectrum comparison.

spectrum as a viable approximation to the frequency response of the seismic data.

Through this approximation, the $U(k)$ becomes a preconditioning operator determined by the dominant frequency. Thus, the new objective function can be formulated as **Equation 16**:

$$E = \int_0^{\frac{\pi}{h}} \left\| U_w\left(k, W\left(f_m,\tau\right)\right)\Phi\left(k\right)\right\|_2^2 dk + \lambda \int_0^{\frac{\pi}{h}} \left\| U_w\left(k, W\left(f_m,\tau\right)\right)\phi\left(k\right)\right\|_2^2 dk \quad (16)$$

where $U_w\,(k, W\,(f_m,\tau))$ depends on the wavenumber and the seismic wavelet, $W\,(f_m,\tau)$ represents the Ricker wavelet generation operator.

The solution to **Equation 16** was found using a least-squares algorithm. According to Liu and Sen,[46] the optimized FD weights obtained by the least-square method can be solved by setting an effective bandwidth ($k_{eff}$). If the objective function value is larger than the setting value, we should decrease the effective bandwidth ($k_{eff}$) to obtain the new FD weights that can output a lower objective function value. Given that the objective function is linear and concave, it has a single lowest point—a global minimum. Thus, we can search for the optimal FD weights using the linear search method. Once the simulation error converges to this minimum, the optimal FD weights for the simulation setup can be determined. These weights result in lower simulation error and allow for a more relaxed stability condition.

This process leverages the key property that the objective function is convex, guaranteeing the existence of a unique global minimum. The search is initialized with the maximum effective bandwidth ($k_{eff} = \pi$) to encompass the full wavenumber range. The algorithm then progressively decrements the effective bandwidth by a small, fixed step size (e.g., $\Delta k = 0.01$). For each step, a set of candidate FD weights is computed by solving a least-squares problem, and the corresponding objective function value is evaluated. By sequentially comparing the objective function values between successive steps, the algorithm precisely identifies the turning point where the function value transitions from decreasing to increasing. An increase in the objective function value relative to the previous step indicates that the search has moved past the global minimum, and therefore the search process is terminated. The effective bandwidth and its associated FD coefficients that yield the minimum function value are then identified as the global optimal solution. This fully automated procedure eliminates heuristic parameter tuning and ensures the robustness and reproducibility of the results.

## 3. Simulation error and stability test

To validate the advantages of the proposed optimized FD weights in improving simulation accuracy, we computed the simulation error with the following simulation parameters: $v$ = 2000 m/s, $\tau$ = 0.001 seconds, h = 10 m, and a dominant frequency $f_m$ = 20 Hz. Based on **Equation 16**, we compared the theoretical simulation errors between the new FD weights and the conventional optimized FD weights. The results, shown in Figure 3, clearly demonstrate a significant reduction in theoretical simulation error achieved by the new FD weights. This indicates that the proposed method can effectively reduce the simulation error.

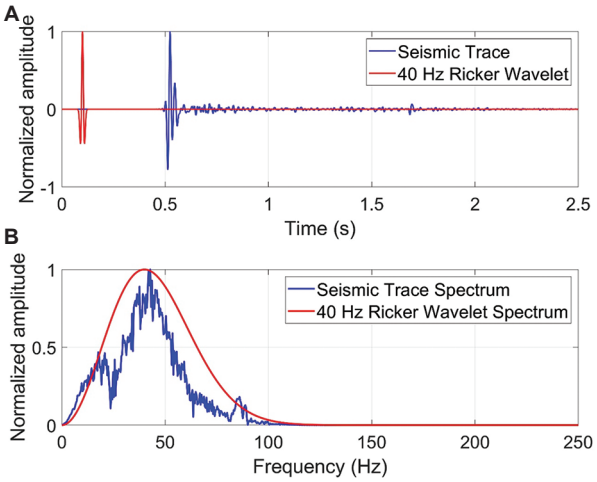



**Figure 2.** Comparative analysis of 40 Hz seismic data and theoretical Ricker wavelet in time and frequency domains. (A) Normalized time–domain comparison. (B) Normalized spectrum comparison.

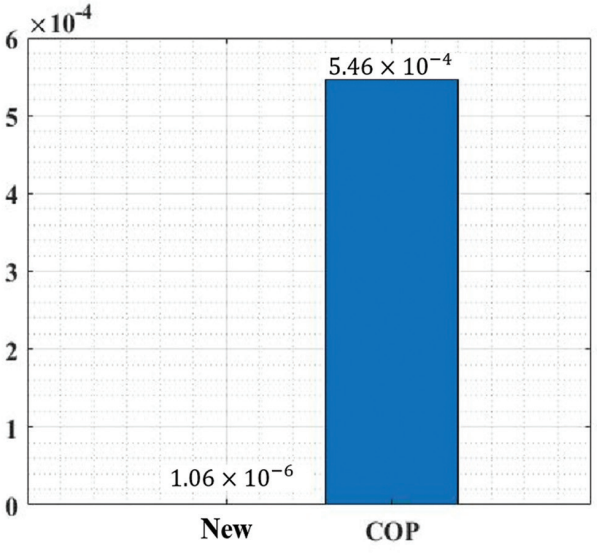



**Figure 3.** The comparison of simulation error $J$ for the new method and the conventional optimized finite-difference weights (COP)

Furthermore, we conducted a comparative analysis from the perspective of numerical stability. For the two-dimensional acoustic wave equations, the stability limit condition can be expressed as **Equation 17**:

$$r_{\max} = \left( 2 \sum_{m=1}^{\mathrm{int}\left((M+1)/2\right)} c_{2m-1} \right)^{-1/2} \tag{17}$$

To evaluate the improvement in stability of the new method, we computed optimized weights using the proposed method within 12th-order FD weights (M=6). We compared the maximum CFL number, ensuring numerical stability for conventional optimized FD weights and the new FD weights. The results, presented in Figure 4, indicate that the new FD weights significantly extended the stability boundary, effectively enhancing the stability of the FD simulation scheme. Two theoretical tests confirmed that the proposed method offers significant advantages in both reducing simulation error and enhancing FD stability.

To evaluate the maximum stable CFL number achievable by the proposed method for M = 3, 4, 5, and 6, the objective function was performed by using the theoretical maximum CFL number—i.e., a value asymptotically approaching unity (since the CFL number must not exceed 1)—rather than the CFL number derived from simulation parameters. Under this limiting condition, the optimized FD coefficients computed by the proposed method were found to remain stable only up to a CFL number of 0.68. Numerical simulations became unstable when the CFL number determined by the simulation parameters exceeded this threshold. Nevertheless, as summarized in Table 1, the proposed method demonstrated a measurable improvement in the maximum allowable CFL number compared with conventional schemes, achieving an average increase of approximately 20%. To validate this stability limit in practice, a numerical simulation was conducted using the optimal FD weights at CFL=0.68. The resulting shot record, presented in Figure 5, was successfully generated without instability, confirming the robustness of our method under this enhanced CFL condition.

**Table 1. Comparison of maximum stable Courant–Friedrichs–Lewy numbers between conventional and optimized finite-difference schemes**

| Operator lengths | 3 | 4 | 5 | 6 |
|---|---|---|---|---|
| Conventional finite-difference weights | 0.575224 | 0.554632 | 0.541266 | 0.531759 |
| Optimal finite-difference weights | 0.680336 | 0.688629 | 0.677551 | 0.684886 |
| Stability improvement | 15.45% | 19.46% | 20.11% | 22.36% |

## 4. FWI imaging test

The objective function of FWI is defined as follows in **Equation 18**:

$$J = \left\| d_{cal} - d_{obs} \right\|_2^2 \tag{18}$$

where $d_{cal}$ represents the simulated data. During the FWI process, $d_{cal} = S_u$, where $S$ denotes the sampling operator, and $u$ is the wavefield, which is the numerical solution of the wave equation. $d_{obs}$ is the observed data.

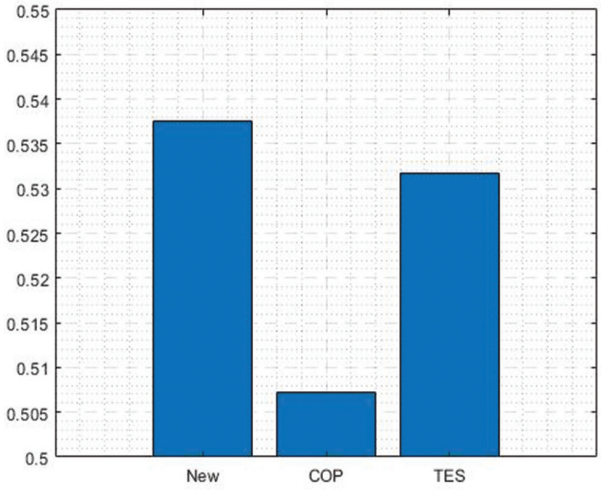



**Figure 4.** Stability comparison for the new method, the conventional optimized finite-difference weights (COP), and the conventional Taylor-expansion method (TES)

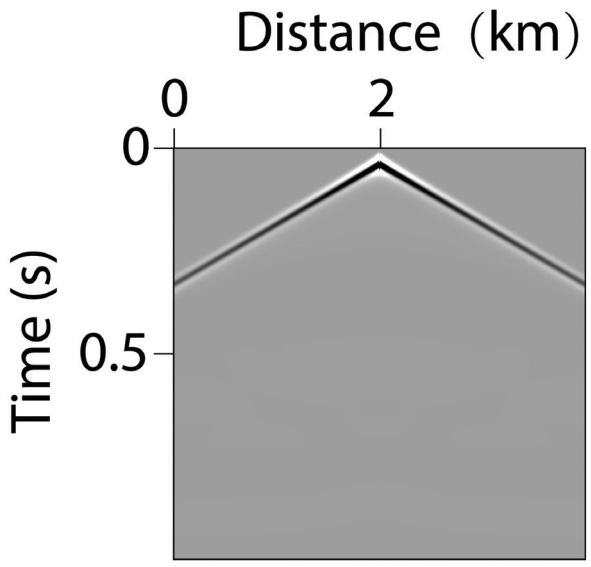



**Figure 5.** A shot record simulated using the optimized finite-difference weights at a Courant–Friedrichs–Lewy number of 0.68

This study focused on the impact of the FD forward-modeling algorithm on FWI imaging results. Here, "FWI imaging result" specifically refers to the spatial derivative of the inverted velocity model, as shown in **Equation 19**:

$$I = \nabla^2 v_{fin} \tag{19}$$

where $I$ denotes the FWI imaging result, $\nabla^2$ represents the second-order spatial derivative, and $v_{fin}$ is the high-resolution full-waveform velocity inversion result. High-resolution FWI requires matching the high-frequency components in the observed data. Due to the nature of FD methods, simulating high frequencies can introduce significant numerical dispersion.

To investigate the impact of numerical dispersion on FWI imaging, a set of numerical experiments was conducted. A horizontal layered model, depicted in Figure 6, was employed as the true velocity model. The model spans 100 grid points in the lateral direction and 600 grid points in depth, with a uniform spatial sampling interval of 10 m. A Ricker wavelet with a dominant frequency of 40 Hz was used as the source function. Wavefield simulation was performed over a total duration of 2 seconds using a time step of 1 ms. The resulting observed record, generated in the absence of a direct wave, is presented in Figure 7A.

Accurate wavefield simulation relies not only on high-precision spatial discretization methods but also on effective boundary treatments to suppress unphysical boundary reflections. To achieve superior absorption, the perfectly matched layer technique was introduced and has since evolved. A perfectly matched layer introduces a special absorbing layer surrounding the computational domain, within which the wavefield is split and attenuated exponentially, thereby efficiently absorbing outgoing waves. Recent improvements, such as the residual perfectly matched layer, have further enhanced absorption performance and numerical stability.[15] In the FWI imaging conducted in this study, to ensure effective control of boundary reflections, the approximation absorbing boundary condition was employed in all our numerical experiments. Our optimized operator was responsible for computing the spatial derivatives across the entire computational domain (including both the interior and the boundary layers), while the absorbing boundary condition acted as a standard module applied at the boundaries to absorb outgoing wave energy.

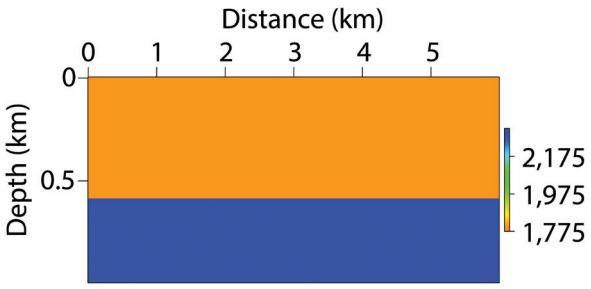



**Figure 6.** The layer model

We conducted wavefield simulation and inversion tests using three different methods to compare the effects of numerical dispersion: (i) A dispersion-free method, which serves as an ideal reference benchmark; (ii) a conventional FD method employing conventional 12$^{th}$-order Taylor-series-based FD weights; and (iii) the proposed new method, in which the weights are optimized using our new scheme.

Figures 7B and 7C show the shot gathers generated from the conventional and new methods, respectively. As shown in Figure 7D, the wavefield simulation results under a dominant frequency of 40 Hz demonstrate that the new method produced only minor numerical dispersion (black curve), while the conventional method exhibits severe dispersive artifacts (blue curve).

Using the simulated data, we performed both objective function analysis and imaging tests within the FWI framework. Due to strong numerical dispersion, the conventional method yielded an objective function (Figure 8A) that contains numerous prominent local minima, and prevented the optimization process from converging to the global minimum. We further studied how numerical dispersion hinders convergence in

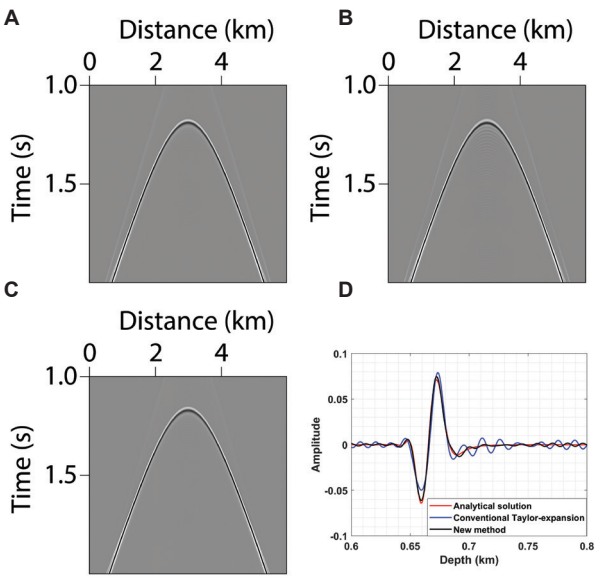



**Figure 7.** Shot gathers comparison for a source located at the middle point of the model. (A) Analytical solution. (B) Conventional Taylor-expansion method. (C) New method (Proposed). (D) Single trace comparison for three solutions.

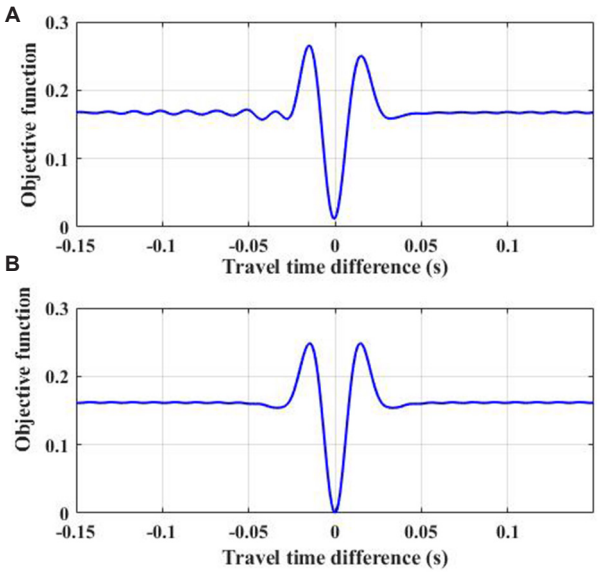



**Figure 8.** Objective function landscapes for velocity estimation. (A) With dispersion error. (B) Without dispersion error (including travel time difference curve).

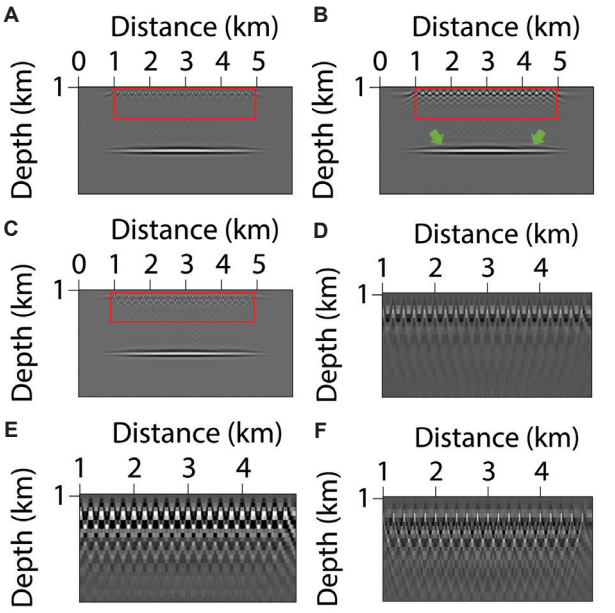



**Figure 9.** FWI velocity model comparison. (A) Reference solution without dispersion. (B) Result from the conventional method with artifacts. (C) Result from the proposed optimized method. (D–F) Zoomed-in views of the shallow target area corresponding to (A–C).

waveform inversion. Dispersion retarded the high-frequency components of a seismic wave, which increased the proportion of low-frequency energy within the main lobe of the arrival, thereby broadening it. As a result, the main lobe becomes broader. In the objective function, this appeared as a wider and smoother central region, which lowered the resolution and accuracy of the model updates.

In contrast, the objective function derived from the dispersion-free simulations (Figure 8B) had more stable convergence behavior. This result demonstrates that numerical dispersion significantly increases the risk of becoming trapped in local minima, thereby affecting both the convergence behavior and the convergence speed of the inversion.

Figure 9 presents the FWI velocity model comparison. Figure 9A shows the dispersion-free reference inversion result, Figure 9B corresponds to the result obtained using the conventional method with strong numerical dispersion, and Figure 9C displays the result from the proposed optimized method. In terms of imaging results, the proposed new method (Figure 9C) closely matched the dispersion-free benchmark (Figure 9A). The events were continuous and clear, with no false structures caused by numerical dispersion. However, the result from the conventional method with strong dispersion (Figure 9B) showed clear artifacts.

These artifacts have distinct spatial patterns. First, lower seismic velocities resulted in higher wavenumbers at a given frequency. Thus, the strong numerical dispersion appeared in the upper part of the model, where velocities were low. This strong dispersion amplified noise related to direct waves—often manifesting as repeated acquisition footprints—making such noise prominent in the shallow section (red box in Figure 9B). Second, the numerical dispersion distorted wavefield paths and imaging geometry, resulting in stronger migration swing artifacts. In the image, these artifacts appeared as strong events below the source location. By comparing Figure 9D-F, we observed that these events had higher energy and covered a larger area, which affected the identification of useful shallow reflections.

To further study how dispersion influences the results, we compared amplitude values from single traces (Figure 9C). The conventional method result (red curve) showed significant deviation from the dispersion-free result (blue curve). In the shallow part, the conventional result showed large amplitude errors. This is mainly because severe dispersion causes a mismatch between the simulated and observed direct waves, leading to incorrect inversion updates near the surface (the black box). In deeper layers, numerical dispersion resulted in mispositioned reflections and oscillatory artifacts (green arrows in Figures 9B and 10). The trace analysis matches the artifacts seen in the full image, confirming that numerical dispersion significantly influenced FWI

results. It creates false reflections, strengthens acquisition footprints and migration swings, and greatly reduces the quality and reliability of the inversion.

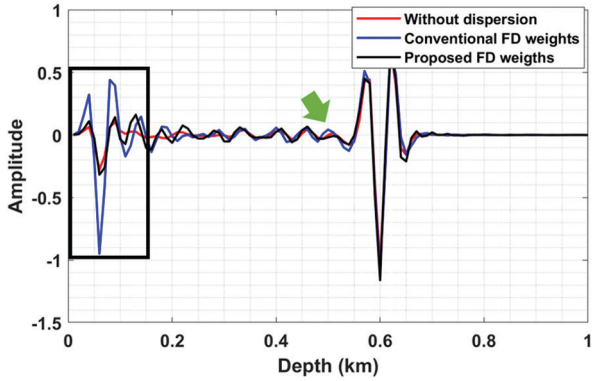



**Figure 10.** The single trace comparison of the full-waveform immersion imaging result in Figure 9
Abbreviation: FD: Finite-difference.

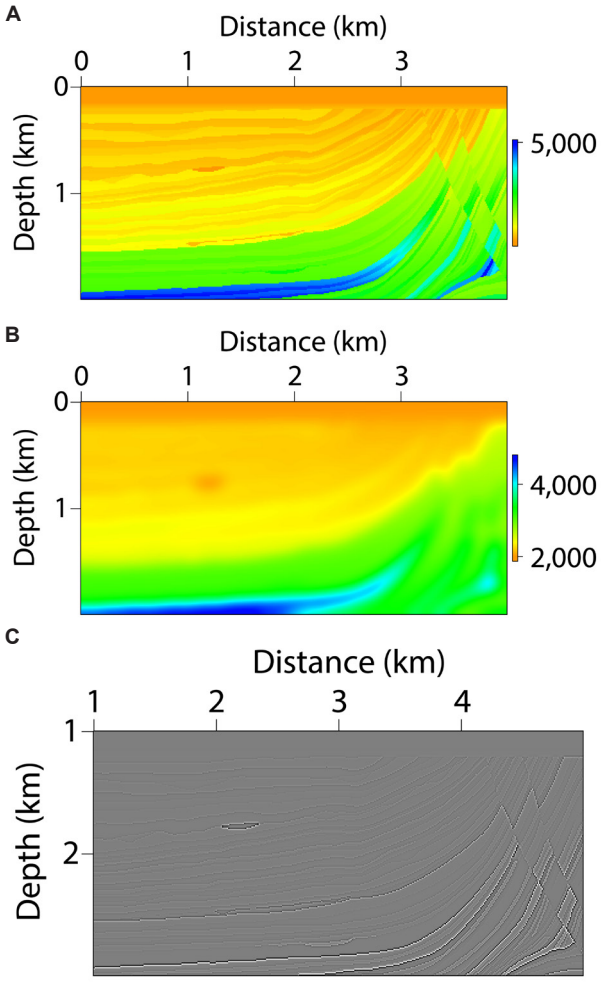



**Figure 11.** The subsurface models. (A) True velocity model. (B) Initial velocity model. (C) True reflectivity model.

To thoroughly evaluate the adaptability of the new method to complex models, we conducted FWI imaging tests using a section of the classic Marmousi model. The model had dimensions of 200 × 400 grid points with a spacing of 10 m between adjacent points. The true velocity model is shown in Figure 11A. The inversion started from a smoothed version of the velocity model (Figure 11B) to reduce the non-linearity of the inverse problem. In the numerical simulation, we used a Ricker wavelet with a dominant frequency of 30 Hz as the source and simulated 2.5 seconds of recorded data with a sampling interval of 1 ms. The acquisition system was configured with 20 shots evenly spaced at 200 m intervals on the surface, and each shot was recorded with full receiver coverage. A sample shot record located at the center of the model is shown in Figure 12.

We first performed comparative FWI experiments using the 30 Hz dominant frequency wavelet. The results indicated that at this frequency, the inversion results from the conventional, new, and dispersion-free reference methods closely matched (Figure 13A-C). The inverted velocity models converged to comparable levels. This indicates that at 30 Hz, numerical dispersion was not significant, and all three modeling methods provided forward wavefields that met the accuracy requirements for inversion. Thus, the objective function decreased stably and converged to a solution near the global minimum.

To further investigate the effect of dispersion, we repeated the inversion using a Ricker wavelet with a dominant frequency of 40 Hz. Under these conditions, the performance of the different methods varied significantly. The conventional method produced shot records with strong numerical dispersion (Figure 14B), particularly in the shallow low-velocity zone, as indicated by the red arrows.

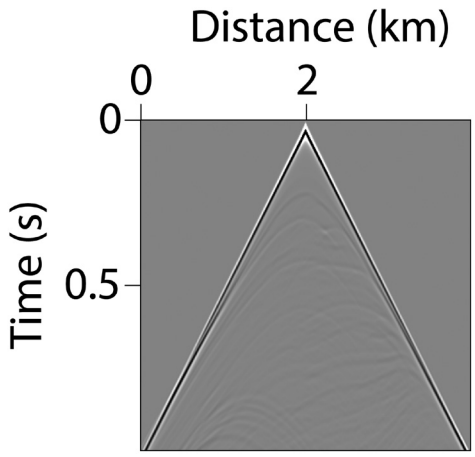



**Figure 12.** The observed data when the dominant frequency of the Ricker wavelet was 30 Hz

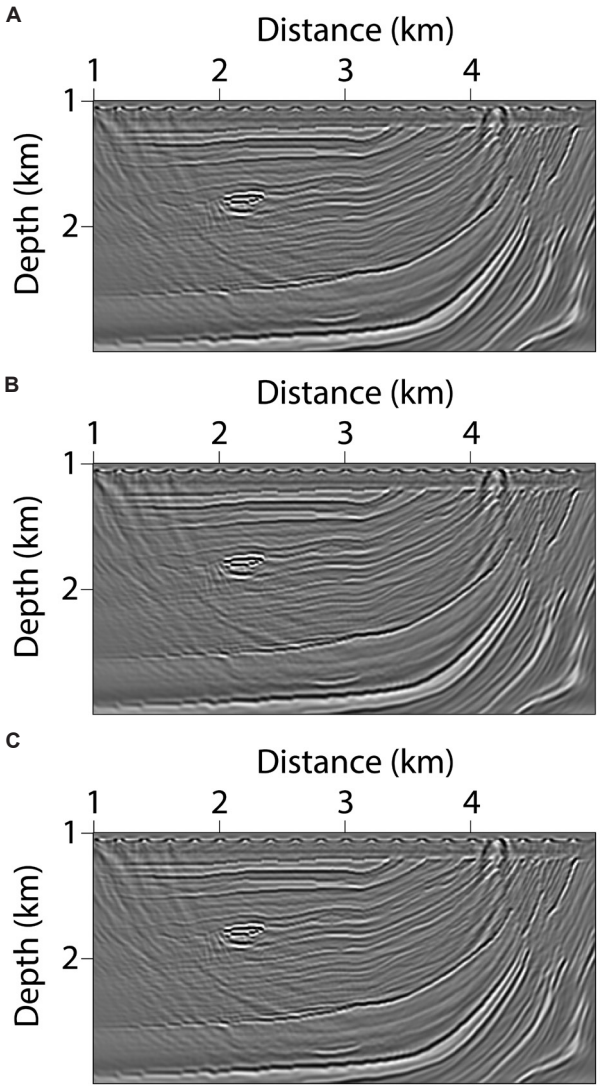



**Figure 13.** Full-waveform inversion imaging results at 30 Hz dominant frequency. (A) Result from the analytical method. (B) Result from the conventional Taylor-expansion method. (C) Result from the new proposed method.

Severe wavefield distortion negatively impacted the inversion process in several ways. First, because dispersion was strongest in the shallow low-velocity zone, wave energy was largely trapped near the surface. This significantly reduced illumination in mid-deep regions, leading to highly unbalanced inversion gradients with depth. Compared with the dispersion-free reference inversion result (Figure 15A), insufficient illumination of deep structures resulted in very weak updates, causing the conventional method to fail to recover the true reflections (as marked by green arrows in Figure 15B). This also caused the inversion gradient to deviate considerably from the true descent direction. Second, severe phase distortion and energy scattering in the waveform's main lobe reduced spatial resolution. Even in shallow areas where updates were relatively concentrated, the recovered events did not achieve the high resolution of that achieved under dispersion-free conditions (as marked by the red arrows in Figure 14B). Ultimately, the conventional inversion process terminated prematurely due to rapidly deteriorating gradient quality, preventing effective reconstruction of deep structures.

In contrast, the new method and the dispersion-free reference were free from these issues. Since their wavefield simulations were unaffected by dispersion, both methods produced well-distributed inversion results and correct phase characteristics. As a result, both methods yielded inversion results with higher resolution and more accurate deep structures (Figure 15A and C). These experiments demonstrated that as the dominant frequency increases and the effect of low-velocity zones becomes more pronounced, numerical dispersion has a greater impact on FWI imaging results. It not only distorts the wavefield phase but also degrades the accuracy of inversion gradients, reducing update efficiency and potentially causing inversion failure. The proposed new method effectively suppressed numerical dispersion at 40 Hz and produced results similar to the dispersion-free reference (Figure 15A and D), confirming its robustness and adaptability in complex models and under higher-frequency conditions.

## 5. Discussion

In this study, the wavefield spectrum $U(k)$ was approximated using the frequency spectrum of the source wavelet $W(k)$ as a practical preconditioner in the optimization problem. While this approximation is computationally efficient and works well in many scenarios, it is important to recognize its limitations, particularly in strongly heterogeneous or highly scattering media. In such environments, wavefield propagation induces frequency shifts, scattering effects, and spectral broadening that are not captured by the source spectrum alone. Consequently, the approximated $U(k)$ may deviate from the true wavefield spectrum. As a result, the FD weights optimized under this approximation can deviate from the globally optimal solution. However, this deviation leads only to sub-optimal, not invalid, FD coefficients.

Although the resulting FD coefficients may be sub-optimal due to the use of an approximate wavefield spectrum, they are still highly effective for practical wavefield modeling. In fact, the so-called sub-optimal coefficients often produce significantly more accurate wavefields than those obtained from conventional FD

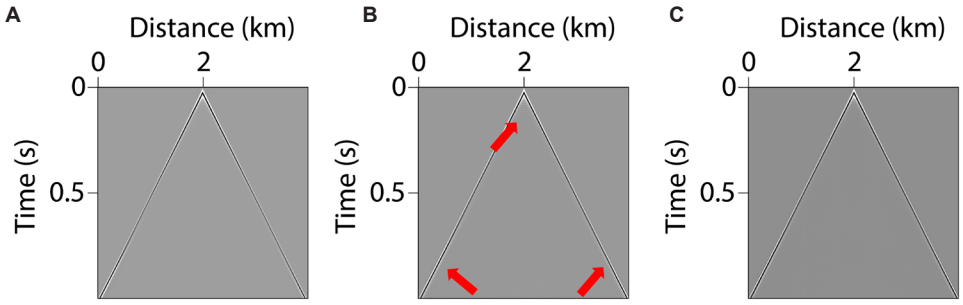



**Figure 14.** Shot gathers comparison for a 40 Hz source. (A) Gather from the analytical method. (B) Gather from the conventional Taylor-expansion method. (C) Gather from the new proposed method.

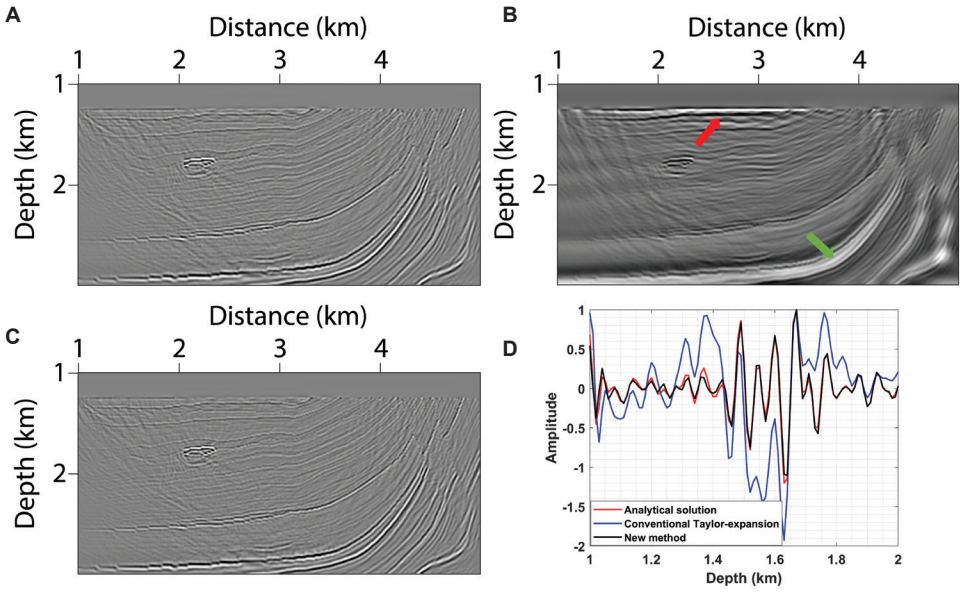



**Figure 15.** Full-waveform inversion imaging comparison at 40 Hz dominant frequency. (A) Analytical method result. (B) Conventional Taylor-expansion method result. (C) New proposed method result. (D) Amplitude profile comparison at a specific location.

methods. This observation is confirmed in our Marmousi model experiment, where the wavefield simulated using the proposed optimized coefficients exhibited reduced numerical dispersion and better fidelity to the true wavefield, ultimately improving the quality of FWI imaging. Therefore, despite the inherent approximation, the method remains robust and beneficial for realistic seismic applications.

Therefore, we regard the spectral approximation not as a limitation that introduces numerical instability, but as a practical and controlled compromise. It allows for computationally efficient optimization of FD weights while delivering robust performance across diverse geological settings. Although the resulting coefficients may be mathematically sub-optimal in a strict sense, they are practically highly effective—consistently reducing numerical dispersion and improving wavefield fidelity, as demonstrated in our Marmousi experiment. This approach thus remains a reliable and beneficial strategy for enhancing the accuracy and quality of FWI in realistic seismic imaging.

## 6. Conclusion

The core of this study is the successful development and validation of a novel method for optimizing FD weights. This method constructed an adaptive objective function that integrates errors from both the time–space and spatial domains, using the seismic wavelet spectrum as a preconditioner. This approach effectively resolved a common industry challenge where conventional methods force a trade-off between simulation accuracy and numerical stability. Theoretically and numerically, we demonstrated that our optimized weights significantly reduced simulation errors across a broad wavenumber

range while substantially relaxing the stability limit, thus permitting larger time steps and greater computational efficiency. Applied to FWI imaging, the method suppressed numerical dispersion, mitigated cycle-skipping issues, and yielded clearer, deeper structural images free from artifacts—even at high frequencies and in complex models—closely approaching the quality of ideal dispersion-free simulations. This work confirms that high-quality forward modeling is crucial for successful FWI and shows that the proposed optimizer extends the performance limits of the FD method, offering strong practical value for advancing high-frequency, high-resolution FWI in real-world applications.

## Acknowledgments

None.

## Funding

This work was supported by the research project "Geophysical Technology and Field Trials for Ultra-Deep Complex Oil and Gas Reservoirs: Research on Q Least-Squares Migration Method" (Grant no. YF202401).

## Conflict of interest

The authors declare that they have no competing interests.

## Author contributions

*Conceptualization:* Ganglin Lei, Jianping Huang
*Data curation:* Zhenwen Liu, Chang Zhou
*Formal analysis:* Weiting Peng
*Funding acquisition:* Jianping Huang
*Investigation:* Jianping Huang, Weiting Peng
*Methodology:* Ganglin Lei, Wensheng Duan, Chao Chen
*Resources:* Jianping Huang
*Writing–original draft:* Ganglin Lei, Wensheng Duan
*Writing–review & editing:* Jianping Huang, Chao Chen, Zhenwen Liu, Chang Zhou, Weiting Peng

## Availability of data

The datasets generated and analyzed during the current study are not publicly available because they contain proprietary information as part of a collaborative project with China National Petroleum Corporation (CNPC). Data sharing is restricted by the confidentiality agreement governing this project.

## References

1. Lailly P. The Seismic Inverse Problem as a Sequence of Before Stack Migrations. In: Bednar JB, editor. *Conference on Inverse Scattering: Theory and Application*. Society for Industrial and Applied Mathematics; 1983. p. 206-220.

2. Tarantola A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*. 1984;49(8):1259-1266.
doi: 10.1190/1.1441754

3. Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics. *Geophysics*. 2009;74(6):WCC1-WCC26.
doi: 10.1190/1.3238367

4. Pratt RG. Seismic waveform inversion in the frequency domain; Part 1, theory and verification in a physical scale model. *Geophysics*. 1999;64(3):888-901.
doi: 10.1190/1.1444597

5. Brossier R, Operto S, Virieux J. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*. 2009;74(6):WCC105-WCC118.
doi: 10.1190/1.3215771

6. Ma X, Xu S, Ke P, Zhang H. Inexact line search method in full waveform inversion. *Appl Geophys*. 2023;20:374-384.
doi: 10.1007/s11770-020-0875-x

7. Warner M, Ratcliffe A, Nangoo T, *et al.* Anisotropic 3D full-waveform inversion. *Geophysics*. 2013;78(2):R59-R80.
doi: 10.1190/geo2012-0338.1

8. Zhang Z, Wu Z, Wei Z, Mei J, Huang R, Wang P. FWI Imaging: Full-Wavefield Imaging through Full-Waveform Inversion. In: *SEG Technical Program Expanded Abstracts 2020*. Society of Exploration Geophysicists; 2020. p. 656-660.
doi: 10.1190/segam2020-3427858.1

9. Lu R. Revealing Overburden and Reservoir Complexity with High-Resolution FWI. In: *SEG Technical Program Expanded Abstracts 2016*. Society of Exploration Geophysicists; 2016. p. 1242-1246.
doi: 10.1190/segam2016-13872562.1

10. Huang R, Zhang Z, Wu Z, Wei Z, Mei J, Wang P. Full-waveform inversion for full-wavefield imaging: Decades in the making. *Lead Edge*. 2021;40(5):324-334.
doi: 10.1190/tle40050324.1

11. Operto S, Gholami Y, Prieux V, *et al.* A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice. *Lead Edge*. 2013;32(9):1040-1054.
doi: 10.1190/tle32091040.1

12. Xu S, Wang D, Chen F, Zhang Y, Lambare G. Full Waveform Inversion for Reflected Seismic Data. In: *Proceedings of 74th EAGE Conference and Exhibition incorporating EUROPEC 2012*. Copenhagen, Denmark: EAGE Publications BV; 2012.
doi: 10.3997/2214-4609.20148725

13. Dong S, Dong X, Zhang R, Cong Z, Zhong T, Wang H. Global-feature-fusion and multiscale network for low-frequency extrapolation. In: *IEEE Transactions on Geoscience and Remote Sensing*. Vol. 62. New York: IEEE; 2024. p. 1-14.

14. Operto S, Gholami A, Aghamiry H, *et al.* Extending the search space of full-waveform inversion beyond the single-scattering Born approximation: A tutorial review. *Geophysics*. 2023;88(6):R671-R702.

doi: 10.1190/geo2022-0758.1

15. Luo Y, Wang T, Li Y, Cai J, Wang Y, Fang G. A residual perfectly matched layer for wave propagation in elastic media. *Acta Geophys*. 2024;72:1561-1573.

doi: 10.1007/s11600-023-01145-x

16. Alterman Z, Karal FC Jr. Propagation of elastic waves in layered media by finite difference methods. *Bull Seismol Soc Am*. 1969;59(1):471.

doi: 10.1785/bssa0590010471

17. Kelly KR, Ward RW, Treitel S, Alford RM. Synthetic seismograms: A finite-difference approach. *Geophysics*. 1976;41(1):2-27.

doi: 10.1190/1.1440605

18. Alford RM, Kelly KR, Boore DM. Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics*. 1974;39(6):834-842.

doi: 10.1190/1.1440470

19. Dablain MA. The application of high-order differencing to the scalar wave equation. *Geophysics*. 1986;51(1):54-66.

doi: 10.1190/1.1442040

20. Holberg O. Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena. *Geophys Prospect*. 1987;35(6):629-655.

doi: 10.1111/j.1365-2478.1987.tb00841.x

21. Kindelan M, Kamel A, Sguazzero P. On the construction and efficiency of staggered numerical differentiators for the wave equation. *Geophysics*. 1990;55(1):107-110.

doi: 10.1190/1.1442763

22. Zhou B, Greenhalgh SA. Seismic scalar wave equation modeling by a convolutional differentiator. *Bull Seismol Soc Am*. 1992;82(1):289-303.

doi: 10.1785/bssa0820010289

23. Zhang JH, Yao ZX. Optimized explicit finite-difference schemes for spatial derivatives using maximum norm. *J Comput Phys*. 2013;250:511-526.

doi: 10.1016/j.jcp.2013.04.029

24. Zhang JH, Yao ZX. Optimized finite-difference operator for broadband seismic wave modeling. *Geophysics*. 2013;78(1):A13-A18.

doi: 10.1190/geo2012-0277.1

25. Yang L, Yan HY, Liu HX. An optimal implicit staggered-grid finite-difference scheme based on the modified Taylor-series expansion with minimax approximation method for elastic modeling. *J Appl Geophys*. 2017;138:161-171.

doi: 10.1016/j.jappgeo.2017.01.020

26. Yang L, Yan HY, Liu H. Optimal staggered-grid finite-difference schemes based on the minimax approximation method with the Remez algorithm. *Geophysics*. 2017;82(1):T27-T42.

doi: 10.1190/geo2016-0171.1

27. He Z, Zhang JH, Yao ZX. Determining the optimal coefficients of the explicit finite-difference scheme using the Remez exchange algorithm. *Geophysics*. 2019;84(3):S137-S147.

doi: 10.1190/geo2018-0446.1

28. Koene EFM, Robertson JOA. Optimal finite-difference operators for arbitrarily sampled data. *Geophysics*. 2020;85(3):F39-F51.

doi: 10.1190/geo2019-0081.1

29. Liu Y. Maximizing the CFL number of stable time-space domain explicit finite-difference modeling. *J Comput Phys*. 2020;416:109501.

doi: 10.1016/j.jcp.2020.109501

30. Yuan W, Rongsheng W. An optimal spatial finite-difference operator which reduces truncation error to a minimum. *Adv Atmos Sci*. 2002;19(3):468-486.

doi: 10.1007/s00376-002-0080-2

31. Ergen JT. A tutorial on optimizing time domain finite-difference schemes: "Beyond Holberg". *Stanford Explo Project*. 2007;129:33-43.

32. Liu Y. Globally optimal finite-difference schemes based on least squares. *Geophysics*. 2013;78(4):T113-T132.

doi: 10.1190/geo2012-0480.1

33. Liu Y. Optimal staggered-grid finite-difference schemes based on least-squares for wave equation modelling. *Geophys J Int*. 2014;197(2):1033-1047.

doi: 10.1093/gji/ggu032

34. Wang YF, Liang WQ, Nashed Z, Li X, Liang CH, Yang CC. Seismic modeling by optimizing regularized staggered-grid finite-difference operators using a time-space-domain dispersion-relationship-preserving method. *Geophysics*. 2014;79(5):T277-T283.

doi: 10.1190/geo2014-0078.1

35. Ren ZM, Liu Y. Acoustic and elastic modeling by optimal time-space-domain staggered-grid finite-difference schemes. *Geophysics*. 2015;80(1):T17-T40.

doi: 10.1190/geo2014-0269.1

36. Yong P, Huang JP, Li ZC, *et al*. Optimized equivalent staggered-grid FD method for elastic wave modelling based on plane wave solutions. *Geophys J Int*. 2017;208(2):1157-1172.

doi: 10.1093/gji/ggw447

37. Miao ZZ, Zhang JH. Reducing error accumulation of optimized finite-difference scheme using the minimum norm. *Geophysics*. 2020;85(5):T275-T291.

doi: 10.1190/geo2019-0758.1

38. Igel H, Weber P, Riollet B. Anisotropic wave propagation through finite-difference grids. *Geophysics*. 1995;60(4):1203-1216.

doi: 10.1190/1.1443849

39. Shao Z, Wei CW, Zhao S. DSC time-domain solution of Maxwell's equations. *J Comput Phys*. 2003;189(2):427-453.

doi: 10.1016/s0021-9991(03)00226-2

40. Xiao F, Tang XH, Zhang XJ. Comparison of Taylor finite difference and window finite difference and their application in FDTD. *J Comput Appl Math*. 2006;193(2):516-534.

doi: 10.1016/j.cam.2005.05.030

41. Chu CL, Stoffa PL, Self R. High-Order Rotated Staggered Finite Difference Modeling of 3D Elastic Wave Propagation in General Anisotropic Media. In: *SEG Technical Program Expanded Abstracts*; 2009. p. 291-295.

42. Chu CL, Stoffa PL. Determination of finite-difference weights using scaled binomial windows. *Geophysics*. 2012;77(3):W17-W26.

doi: 10.1190/geo2011-0336.1

43. Remez EY. Sur la determination des polynomes D'approximation de degre donnee [On the determination of approximation polynomials of given degree]. *Commun Kharkov Math Soc*. 1934;10:41-63.

44. McClellan J, Parks T, Rabiner L. A computer program for designing optimum FIR linear phase digital filters. *IEEE Trans Audio Electroacoustics*. 1973;21(6):506-526.

doi: 10.1109/tau.1973.1162525

45. Peng WT, Huang JP. A robust seismic wavefield modeling method based on minimizing spatial simulation error using $L_2$-norm cost function. *Pet Sci*. 2024;22(3):1051-1061.

doi: 10.1016/j.petsci.2024.12.003

46. Liu Y, Sen MK. A new time-space domain high-order finite-difference method for the acoustic wave equation. *J Comput Phys*. 2009;228(23):8779-8806.

doi: 10.1016/j.jcp.2009.08.027

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Elastic modulus analysis of a 3D-printed permafrost simulation model with varying porosity in a seismic physical modeling

**Daechul Kim[1], Yeonjin Choi[1], Wookeen Chung[2], and Sungryul Shin[2]***

[1]Division of Glacier and Earth Sciences, Korea Polar Research Institute, Incheon, Republic of Korea

[2]Department of Energy and Resources Engineering, College of Ocean Science and Engineering, National Korea Maritime and Ocean University, Busan, Republic of Korea

## Abstract

Global warming is causing permafrost thawing, which can lead to permafrost collapse. Assessing the stability of permafrost against this collapse risk is essential. One effective approach to evaluating permafrost stability is to use the elastic modulus. Since the elastic modulus varies with the ice content in the pore spaces, analyzing its relationship with porosity is crucial for understanding permafrost stability. Previous studies have used rock cores to analyze the relationship between elastic modulus and porosity. However, the analysis of elastic modulus relative to various porosities in permafrost is limited. To overcome these limitations, porosity was controlled using 3D printing, and a permafrost analog model with seismic velocities and porosities similar to those of natural permafrost was fabricated and assumed as permafrost. To analyze elastic modulus as a function of porosity in permafrost, seismic velocities were measured through seismic physical modeling experiments, and the elastic modulus was estimated from the measured seismic velocities. Analysis of the relationship between elastic modulus and porosity revealed that the elastic modulus decreased after increasing to a specific porosity in the permafrost simulation model. These findings provide a quantitative basis for evaluating the stability of infrastructure on permafrost by demonstrating the significant effect of ice expansion on mechanical properties. Furthermore, this study validates the 3D printing approach as an effective tool to overcome the limitations of natural rock cores for systematic permafrost research.

***Keywords:*** Seismic physical modeling; Permafrost; 3D printing; Porosity; Elastic modulus

***Corresponding author:**
Sungryul Shin
(srshin@kmou.ac.kr)

**Citation:** Kim D, Choi Y, Chung W, Shin S. Elastic modulus analysis of a 3D-printed permafrost simulation model with varying porosity in a seismic physical modeling. *J Seismic Explor*. 2026;35(2):025420091. doi: 10.36922/JSE025420091

**Received:** October 17, 2025

**Revised:** January 2, 2026

**Accepted:** January 16, 2026

**Published online:** March 6, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Permafrost is ground that remains below 0°C for at least 2 consecutive years, and covers about 15% of the Northern Hemisphere.[1,2] Permafrost stores a large amount of organic carbon, and its melting due to global warming can accelerate global warming by increasing emissions of greenhouse gases such as carbon dioxide and methane.[3,4] Specifically, the thawing of permafrost is causing ground collapse, damaging pipes, infrastructure, etc.[5] Therefore, permafrost analysis is necessary to predict and prepare for changes in permafrost caused by thawing to ensure stability. One method for analyzing permafrost stability is the elastic modulus.[6]

The permafrost pores are saturated with ice, leading to varying ice saturation relative to porosity. Because ice thawing within pores can cause ground collapse, analyzing the elastic modulus as a function of porosity is crucial for assessing and predicting the stability of permafrost under thawing. The elastic modulus of permafrost can be measured directly using destructive testing of core samples or indirectly using seismic velocity. To analyze seismic velocity relative to porosity in permafrost, core samples are measured in the laboratory, or well logging is performed. However, because permafrost varies regionally and is found in polar regions, obtaining samples with diverse porosities or conducting well logging is difficult. Therefore, prior research has used rock cores to simulate permafrost and analyze porosity and seismic velocity.[7-11] Recently, research has been conducted to fabricate frozen soil models using artificial porous sandstone samples and to estimate ice saturation.[12] Although rock core samples and artificial porous sandstone samples have the advantage of being media similar to the permafrost environment, there are limitations in controlling the pore size, structure, and porosity desired by researchers. In contrast, 3D printing enables the control of the porous model as desired by the researcher. If porosity can be controlled and its elastic modulus can be analyzed, this will help assess the stability of permafrost.

As a method to create 3D models by layering, 3D printing offers the advantage of producing porous models with complex structures. It can be classified into seven types based on the additive manufacturing method. These include material extrusion, vat photopolymerization, powder bed fusion, material jetting, binder jetting, direct energy deposition, and sheet lamination.[13] Studies have used 3D printing to create rock cores that simulate pores and analyzed their properties. Ishutov *et al.*[14] proposed a method for simulating the porosity of sandstone using 3D printing with material extrusion and vat photopolymerization. Meanwhile, Ibrahim *et al.*[15] compared and verified the porosity of 3D-printed rock cores with that of real rock cores using vat photopolymerization and material jetting.

Utilizing this 3D printing technique, porous models can be created by controlling porosity, and their properties can be analyzed based on that porosity. Seismic physical modeling is a method for measuring seismic velocity, a property of 3D-printed porous models. This technique utilizes actual propagating waves and elastic media. Although the ultrasonic frequency and model dimensions differ from those in field exploration, this method offers the advantage of analyzing wave propagation characteristics within a controlled laboratory environment. Recently, this method has been effectively applied to analyze the seismic response characteristics of P, shear-vertical (SV), and shear-horizontal (SH) waves to reservoir parameters and to study complex lithofacies and structures in shale oil reservoirs.[16,17]

Studies have utilized 3D printing to create porous models for seismic physical modeling, saturate the pores with fluid, and analyze the fluid. Huang *et al.*[18] and Huang *et al.*[19] analyzed seismic velocity by saturating water in a penny-shaped porous model made by 3D printing using the material extrusion method, while Dande *et al.*[20] analyzed seismic velocity and anisotropy as a function of fluid and viscosity in the same model. Dande *et al.*[21] and Wang *et al.*[22] fabricated porous models with fractures using 3D printing and analyzed the elastic wave characteristics of the fractures. Zerhouni *et al.*[23] and Kim *et al.*[24] used 3D printing to print porous models with varying pore sizes and analyzed the seismic properties with respect to porosity.

In seismic physical modeling, studies of porous models printed in 3D have primarily focused on saturating fluids and analyzing seismic velocities. The present study simulated permafrost by saturating the pores of 3D-printed porous models with varying porosities with ice. Seismic velocities were measured using the simulated permafrost models, and the corresponding elastic moduli were calculated to analyze how elastic modulus changes with porosity. The elastic modulus is one of the important physical properties in determining the stability of permafrost during thawing. The porosity of the permafrost simulation model was designed to fall within the permafrost porosity range, and the medium used was similar to the velocity of the strata that comprised the permafrost.

## 2. 3D-printed permafrost simulated model

### 2.1. Design and printing of a permafrost simulation model with similar porosity and P-wave velocity to permafrost

A permafrost simulation model was designed and 3D-printed. To ensure the model was similar to permafrost, the porosity and P-wave velocity were set to values similar to those reported in previous studies. Figure 1 summarizes previous studies on the porosity and P-wave velocity of permafrost. The porosity ranges from 10% to 60%, and the P-wave velocity was 2,000–6,000 m/s.[7-11] In the present study, permafrost simulation models were designed with porosities of 23.2%, 30.0%, 38.0%, and 42.7% to span the permafrost porosity range. The porosity was calculated as the ratio of the total volume to the volume of pores in the design. The porosity was controlled by varying the number of pores, and the pore size was kept constant at 2 mm. In addition, to facilitate water injection and ensure complete ice saturation within the pores, 0.5 mm

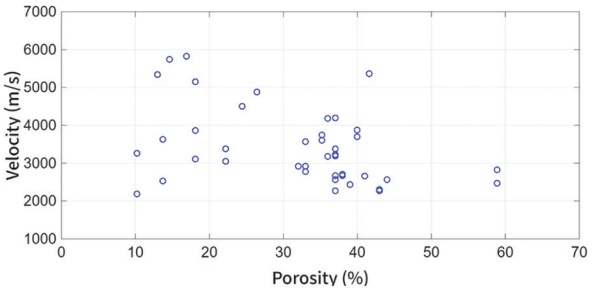



**Figure 1.** Permafrost porosity and P-wave velocity. Blue markers represent the P-wave velocity of permafrost as a function of porosity, as reported in prior research.[7-11]

channels were added to the X and Y faces, each connecting to every pore. Figure 2 presents the schematic design of the permafrost simulation models, detailing the internal structure, which features 2 mm pores and a network of 0.5 mm interconnecting channels on the X and Y faces to ensure effective fluid flow and complete saturation. Figure 2A represents Model 1, designed with a porosity of 23.2%, Figure 2B shows Model 2 with 30.0%, Figure 2C corresponds to Model 3 with 38.0%, and Figure 2D depicts Model 4 with 42.7%.

The designed permafrost simulation models were intended to be printed on a medium with a P-wave velocity similar to that of the permafrost. To measure the P-wave velocity of the polylactic acid (PLA) filament, a PLA solid model with each axis of 25 mm and a porosity of 0% was printed. Its P-wave velocity was measured using seismic physical modeling and was approximately 2,150 m/s.[25] Meanwhile, the velocity of ice-free sedimentary layers in permafrost areas was <2,500 m/s. As the P-wave velocity of the PLA filament was similar to that of the permafrost, the permafrost simulation models were 3D-printed with the PLA filament using fused deposition modeling (FDM) (Figure 3).[26] FDM is a representative additive manufacturing technology based on material extrusion, wherein a thermoplastic filament is heated to its melting point, extruded through a nozzle, and solidified on cooling to form successive layers.

### 2.2. Verification of the quality of the permafrost simulation model using porosity measurements

To verify the quality of the 3D-printed permafrost simulation models, porosities were measured. If it were measured similarly to the design model, it was thought that the porosity calculated in the design drawing and the porosity of the 3D-printed model would be similar. To measure the porosity of the permafrost simulation model, **Equation 1** was used:

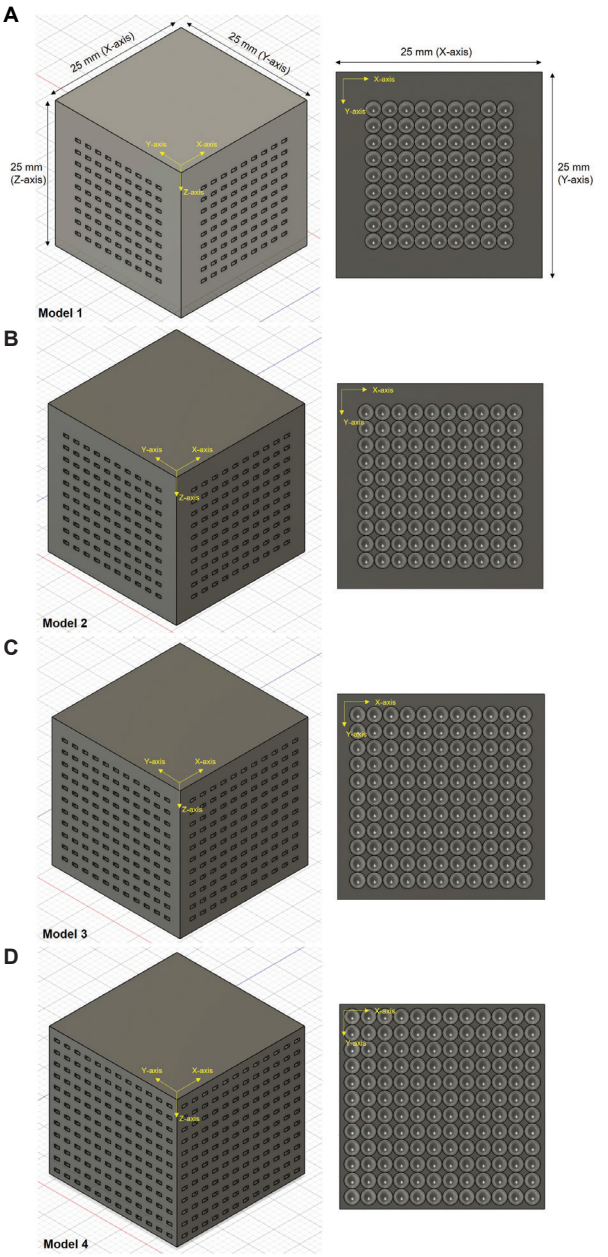



**Figure 2.** Permafrost simulation models designed with various porosities. (A) Model 1. (B) Model 2. (C) Model 3. (D) Model 4.

$$\varnothing = \frac{V_v}{V} \times 100\ (\%) \tag{1}$$

where $\varnothing$ is the porosity, $V$ is the total volume of the model, and $V_v$ is the void volume. $V_v$ was calculated using **Equation 2**, and each parameter is listed in Table 1:

$$V_v = \frac{M - M_M}{\rho} \tag{2}$$

**Table 1. Porosities of the permafrost simulation models and the parameters used for porosity calculation**

| Model | Measured porosity (Ø, %) | Mass without pores ($M$, g) | Mass with pores ($M_M$, g) | Total volume ($V$, cm³) | Void volume ($V_v$, cm³) | Density ($\rho$, g/cm³) |
|---|---|---|---|---|---|---|
| 1 | 23.4 | 19.3 | 14.7 | 15.8 | 3.7 | 1.2 |
| 2 | 30.2 | 19.3 | 13.4 | 15.7 | 4.7 | 1.2 |
| 3 | 38.3 | 19.3 | 11.8 | 15.7 | 6.0 | 1.2 |
| 4 | 47.6 | 19.3 | 10.0 | 15.7 | 7.5 | 1.2 |

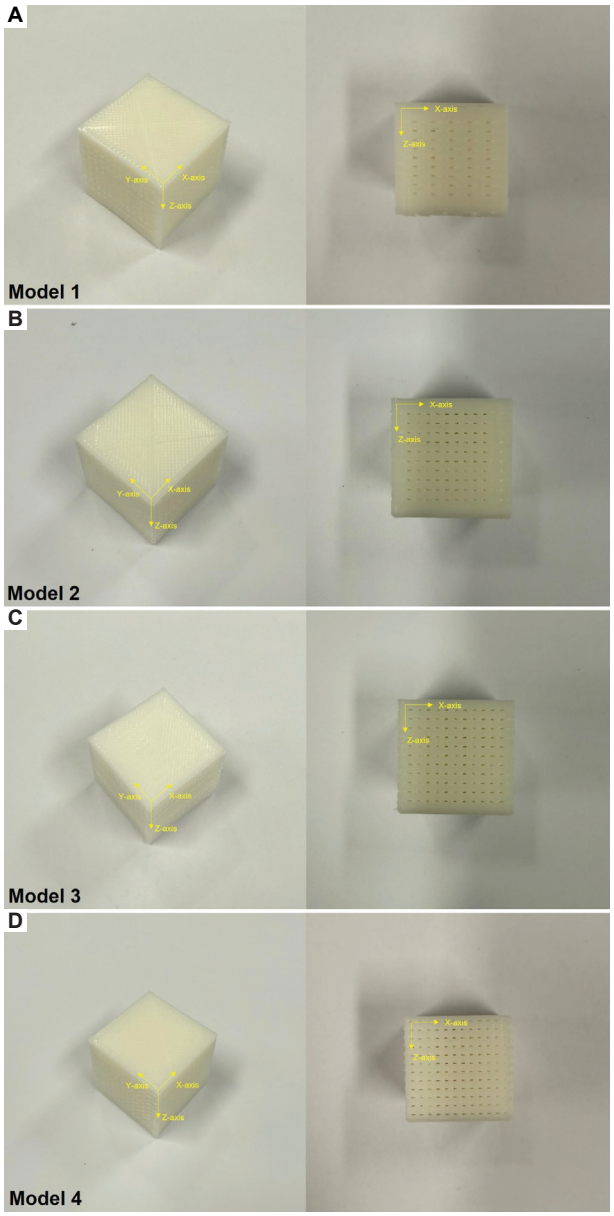



**Figure 3.** 3D-printed permafrost simulation models. (A) Model 1. (B) Model 2. (C) Model 3. (D) Model 4.

where $M$ is the mass of the model without pores, $M_M$ is the mass of the permafrost simulation model with pores, and $\rho$ is the density of the model. The porosity measurements were 23.4%, 30.2%, 38.3%, and 47.6% for Models 1, 2, 3, and 4, respectively. These porosities were comparable to the corresponding calculated porosities, indicating the quality of the permafrost simulation models (Figure 4).

### 2.3. Permafrost simulation

To simulate the characteristics of permafrost within a 3D-printed model, a controlled water saturation process was implemented using a vacuum chamber and a vacuum pump. The purpose of this procedure was to ensure full saturation of the internal pore structures of the 3D-printed sample with water. As shown in Figure 5, the model was first immersed in water within the vacuum chamber. Subsequently, a vacuum pump was used to reduce the ambient pressure, thereby facilitating the infiltration of water into the internal pores by displacing entrapped air.

In this study, it is essential to ensure that the pore space within the permafrost simulation model is fully saturated with ice after freezing. To achieve this condition, the pores must first be completely saturated with water before the freezing process. Therefore, to verify the degree of saturation, porosity measurements were conducted at different time intervals during the saturation process. As illustrated in Figure 6, the measured porosity remained constant regardless of the saturation duration, and the values closely matched the initial porosity obtained through the method described in Section 2.2. This consistency indicates that the internal pores had been effectively and fully saturated with water. Hence, it was concluded that the sample reached full saturation, making it suitable for subsequent freezing and seismic physical modeling.

To simulate the formation of permafrost, the water-saturated porous model was subjected to a controlled freezing process by maintaining it at a temperature of −25°C for a duration of 24 h. During freezing, water within the pore spaces expands on phase transition to ice, which can exert internal stresses and potentially lead to structural deformation or cracking of the 3D-printed model. Therefore, a deformation control strategy was

implemented to prevent structural distortion of the model during the freezing process.

As illustrated in Figure 7, the model was mechanically constrained along the Z-axis using a vice to prevent vertical expansion. This fixation effectively minimized deformation in the vertical direction, where stress accumulation is most likely to occur due to gravity-assisted water distribution. In addition, the model was designed and positioned such that excess water, subjected to volume expansion on freezing, could be expelled laterally along the X- and Y-axes. By allowing this lateral release of pressure, the risk of internal damage and structural distortion within the model was significantly reduced. This approach ensured that the model retained its original shape and dimensions after freezing, thereby maintaining the reliability of the permafrost simulation and subsequent experimental analyses.

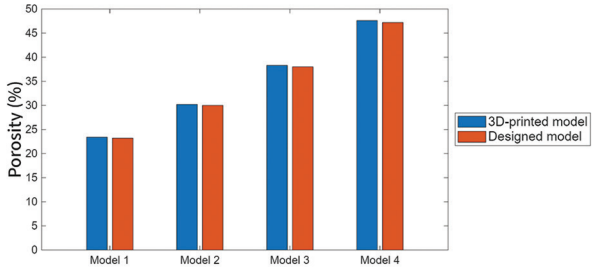



**Figure 4.** Comparison of measured and calculated porosities across permafrost simulation models

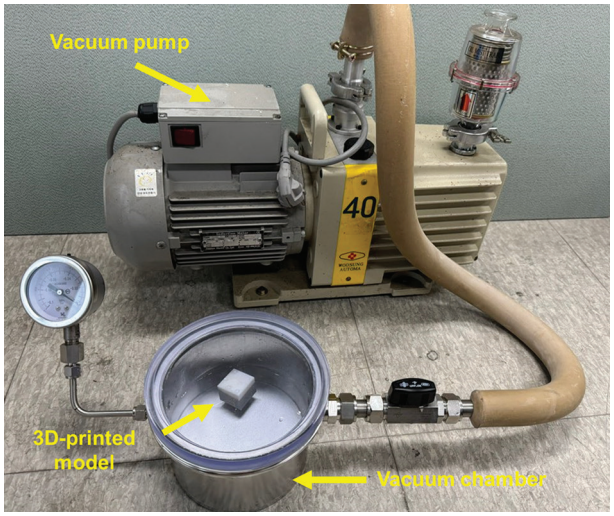



**Figure 5.** A system for saturating the pores of the 3D-printed models

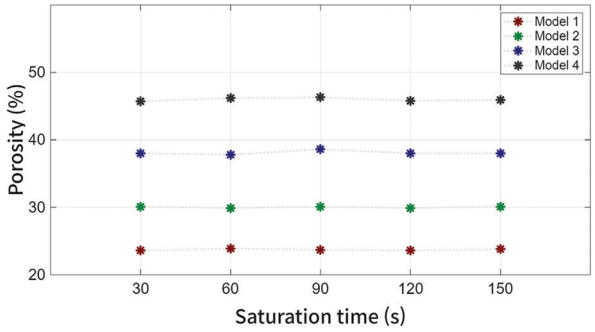



**Figure 6.** Porosity of 3D-printed models over saturation time

To verify that the pores within the 3D-printed permafrost model were fully saturated with ice after freezing, the porosity of the frozen model was measured. The measurement was conducted using the same method described in Section 2.2, and the results confirmed that the porosity was similar to the initial porosity before saturation (Table 2). Therefore, it was determined that the pores within the model were fully saturated with ice. The slight discrepancy in porosity measurements before and after freezing is an inevitable consequence of the freezing expansion of water. Although deformation was mechanically constrained along the Z-axis, the internal expansion pressure of the ice may have induced minor deformations in the PLA skeleton. However, these results confirm that the models were successfully saturated with ice without experiencing significant structural failure. Accordingly, the porosity measured after freezing was used for the research analysis. Further analysis linking these expansion-induced stresses to the observed elastic modulus behavior is provided in the discussion section.

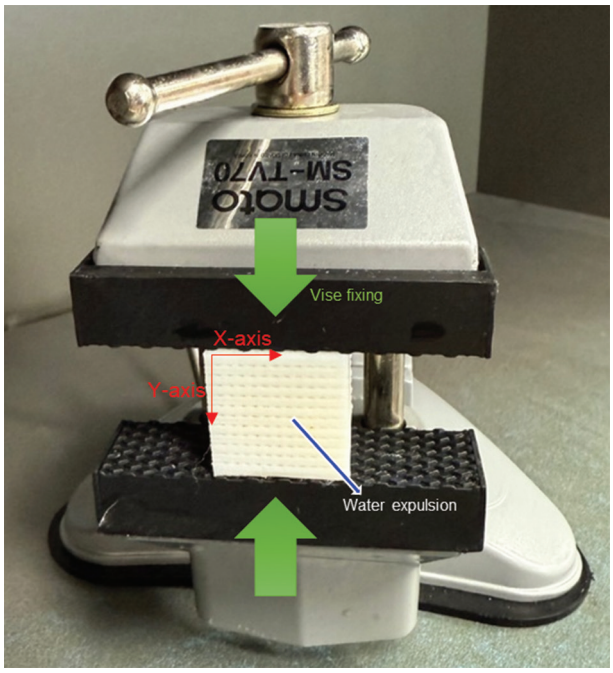



**Figure 7.** Preventing deformation of permafrost simulation models due to volume expansion

## 3. Seismic physical modeling

In this study, seismic physical modeling was conducted to estimate the elastic modulus of permafrost by measuring seismic velocities. This method enables the measurement of P-wave and S-wave velocities within a permafrost simulation model designed to replicate the porosity and seismic velocities of permafrost. A pulser/receiver system (5055PR, Panametrics, USA) was employed to generate and control the source signals, while a digital oscilloscope (MSOX2012A, Keysight, USA) was utilized for data acquisition, enabling precise recording of waveforms transmitted through the model. The overall experimental setup is illustrated schematically in Figure 8.

For wave generation and detection, 1 MHz P-wave and S-wave piezoelectric transducers (V103 and V153, for P-wave and S-wave, respectively, Olympus, USA) were used as the source and receiver. These transducers were carefully aligned along their respective propagation axes and firmly fixed in place to ensure consistent and repeatable signal transmission and reception. S-wave data were obtained by measuring both SH and SV wave components.

The SH-waves were acquired by positioning the S-wave transducer such that the vibration direction was perpendicular to the direction of wave propagation and parallel to the horizontal plane. To acquire SV-waves, the same S-wave transducer was rotated by 90° from the SH-wave orientation, thereby aligning its vibration direction vertically relative to the horizontal propagation path. This configuration allowed the separate characterization of different shear wave modes, which is essential for accurately estimating shear-related elastic moduli within the simulated permafrost model. Furthermore, measurements for each waveform were repeated five times to ensure reproducibility.

Figure 9 shows the P, SH, and SV wave waveforms obtained from the seismic physical modeling, and only the waveforms along the X-axis of Model 1 are shown as representative examples. The noise caused by the trigger at 0 μs was removed by muting.

**Table 2. Comparison of measured porosity before and after freezing to verify ice saturation**

| Model | Measured porosity before freezing (%) | Measured porosity after freezing (%) |
|---|---|---|
| 1 | 23.4 | 23.6 |
| 2 | 30.2 | 30.0 |
| 3 | 38.3 | 37.8 |
| 4 | 47.6 | 45.9 |

## 4. Elastic modulus analysis for porosity in permafrost simulation models

One method for analyzing the stability and condition of permafrost is to use the elastic modulus.[27,28] Therefore, to determine the stability of permafrost with respect to porosity, the elastic modulus was analyzed with respect to porosity. To analyze the elastic modulus for porosity in a permafrost simulation model, seismic velocities were calculated. The seismic velocities were derived from waveforms acquired, and the travel time from 0 μs to the signal and the model length were calculated. Table 3 shows the average P-wave and S-wave velocities, derived from five repeated measurements. The margin of error for these measurements ranged from 0.1% to 0.5%. The presence of elastic anisotropy within the permafrost simulation models can be confirmed from these velocity measurements. In general, the P-wave velocities showed distinct directional dependence, with horizontal velocities (X- and Y-axes) consistently exceeding the vertical velocity

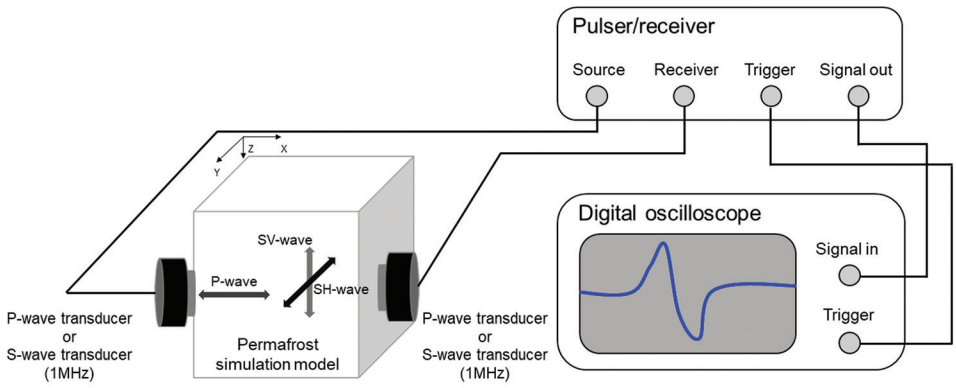



**Figure 8.** The schematic diagram of seismic physical modeling in the permafrost simulation model
Abbreviations: SH: Shear-horizontal; SV: Shear-vertical.

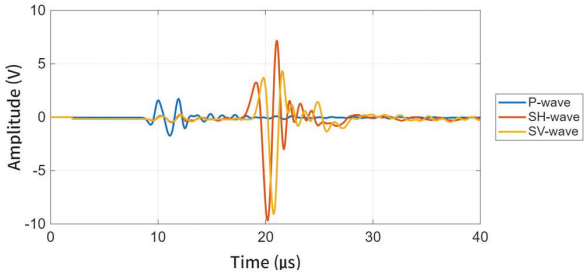



**Figure 9.** Waveforms of P, shear-horizontal (SH), and shear-vertical (SV) waves along the X-axis of Model 1

(Z-axis). Furthermore, significant shear wave splitting was observed, where SH-wave velocities were generally higher than SV-wave velocities. This anisotropy is attributed to the layer-by-layer manufacturing process of 3D printing, and it has been reported in previous studies using 3D-printed models.[19]

To verify the seismic physical modeling results, permafrost P-wave velocity and the Wylie equation for various porosities were compared (Figure 10). The Wyllie equation, a porous rock model for estimating porosity from seismic velocity, was used in this study to verify the velocity measurements with respect to porosity. This is because the measured seismic velocities must be accurate to be used in elastic modulus analysis. The Wyllie equation is shown in **Equation 3**[29]:

$$\frac{1}{V_p} = \frac{\varnothing}{V_f} + \frac{1-\varnothing}{V_m}.V_m \quad (3)$$

Where $V_p$ is the P-wave velocity, $\emptyset$ is the porosity, $V_f$ is the saturated fluid velocity, and $V_m$ is the medium velocity.

To calculate the Wyllie equation, we acquired data from a 25 mm cube PLA solid model, identical to the permafrost simulation model, and seismic physical modeling data for ice (Figure 11). The cube PLA solid model has a 0% porosity and was measured after 24 h of freezing at −25°C, the same conditions as the permafrost simulation model. The velocities were 2,590.8 m/s, 2,578.0 m/s, and 2,552.3 m/s along the X-, Y-, and Z-axes, respectively. The average velocity of the three axes, 2,573.7 m/s, was used as $V_m$ in the Wyllie equation. The ice velocity, 4,090.3 m/s, was measured under the same experimental conditions and was used as $V_f$. The verification results showed that the experimental results are within the velocity range of the permafrost in the seismic physical modeling and are consistent with the calculated results using the Wyllie equation.

The elastic moduli were calculated using the seismic velocities, which were the averages of five repeated

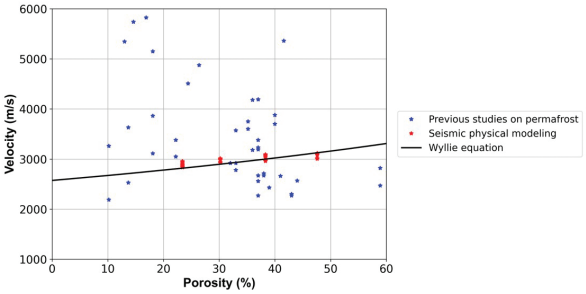



**Figure 10.** Graph demonstrating P-wave velocity against porosity. The permafrost velocities for various porosities shown in blue are referenced from Timur *et al.*,[7] Zimmerman and King,[8] Dou *et al.*,[9] Li and Matsushima,[10] and Kurfurst.[11] The red markers represent the P-wave velocities relative to porosity obtained from seismic physical modeling, while the black line indicates the theoretical results calculated using the Wyllie equation.

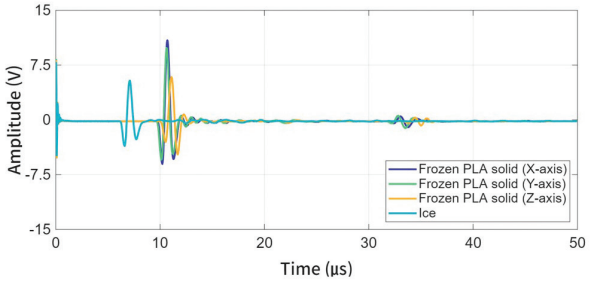



**Figure 11.** Reference P-wave waveform for calculating the Wyllie equation
Abbreviation: PLA: Polyactic acid.

**Table 3. Seismic velocity for elastic modulus analysis in permafrost simulation models**

| Model | Axis | Porosity (%) | P-wave velocity (m/s) | SH-wave velocity (m/s) | SV-wave velocity (m/s) |
|---|---|---|---|---|---|
| 1 | X | 23.6 | 2,942.1 | 1,410.1 | 1,342.2 |
| | Y | | 2,883.9 | 1,399.9 | 1,328.2 |
| | Z | | 2,842.9 | 1,343.0 | 1,330.3 |
| 2 | X | 30.0 | 3,001.7 | 1,431.0 | 1,370.1 |
| | Y | | 2,990.8 | 1,428.0 | 1,356.3 |
| | Z | | 2,928.0 | 1,375.2 | 1,376.2 |
| 3 | X | 37.8 | 3,085.7 | 1,472.3 | 1,414.0 |
| | Y | | 3,043.0 | 1,463.6 | 1,423.3 |
| | Z | | 2,967.2 | 1,409.8 | 1,410.3 |
| 4 | X | 45.9 | 3,110.2 | 1,484.6 | 1,444.7 |
| | Y | | 3,087.7 | 1,481.9 | 1,428.4 |
| | Z | | 3,010.3 | 1,439.6 | 1,428.7 |

measurements. Young's modulus ($E$), shear modulus ($G$), and bulk modulus ($K$) were calculated using **Equations 4-6**.[30]

$$E = V_p^2 \times \rho \quad (4)$$

$$G = V_s^2 \times \rho \quad (5)$$

$$K = \rho(V_p^2 - \frac{4}{3}V_s^2) \quad (6)$$

Where $V_p$ is the P-wave velocity, $\rho$ is the density, and $V_s$ is the S-wave velocity. The resulting error ranges were 0.4–1.0% for Young's modulus, 0.2–0.7% for the shear modulus, and 0.4–1.4% for the bulk modulus.

Figure 12 shows Young's modulus as a function of porosity in a permafrost simulation model. The $R^2$ values displayed alongside the equations represent the coefficient of determination, indicating the goodness of fit for the regression models. As porosity increased, Young's modulus also increased, but then decreased at a porosity of approximately 38%. Figure 13 illustrates the shear modulus as a function of porosity, calculated using SH- (Figure 13A) and SV-wave (Figure 13B) velocity and density data obtained from seismic physical modeling. Shear modulus also increased with increasing porosity, then decreased at approximately 38%.

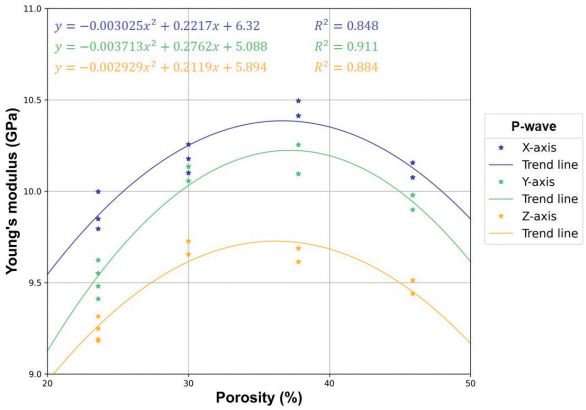



**Figure 12.** Young's modulus versus porosity in the permafrost simulation model. The trend lines were generated using a second-order polynomial. Blue, green, and yellow markers represent the Young's modulus measured along the X-, Y-, and Z-axes, respectively.

Bulk modulus was calculated using P-wave, S-wave, and density. The S-wave was calculated by distinguishing between the SH-wave and the SV-wave. Figure 14 presents the relationship between porosity and bulk modulus, where Figure 14A represents the results derived from SH-waves and Figure 14B represents those from SV-waves. As porosity increased, bulk modulus also increased and then gradually decreased.

In permafrost, the elastic modulus is affected by a combination of factors, including pore size, porosity, temperature, and the degree of ice saturation within the pore space. Among these, porosity and ice content play particularly important roles in determining the mechanical behavior of frozen ground. Previous studies have generally shown that when pores are fully saturated with ice, the elastic modulus tends to increase with increasing porosity, primarily due to the stiffening effect of ice within the pore network.[31,32]

However, in this study, non-linear trends were observed in the relationship between elastic modulus and porosity. Specifically, the elastic modulus increased with increasing porosity up to a certain threshold, beyond which it began to decrease. To understand this behavior, the analysis focused on how seismic velocity and density vary with porosity, along with the effects of freezing on these properties. This is because the elastic modulus was calculated using **Equations 4-6**, which are functions of both seismic wave velocity and material density.

Figure 15 presents the variation of both parameters as porosity increased under fully ice-saturated conditions. Since P-wave, SH-wave, and SV-wave velocities were utilized to calculate the elastic modulus, the variations in

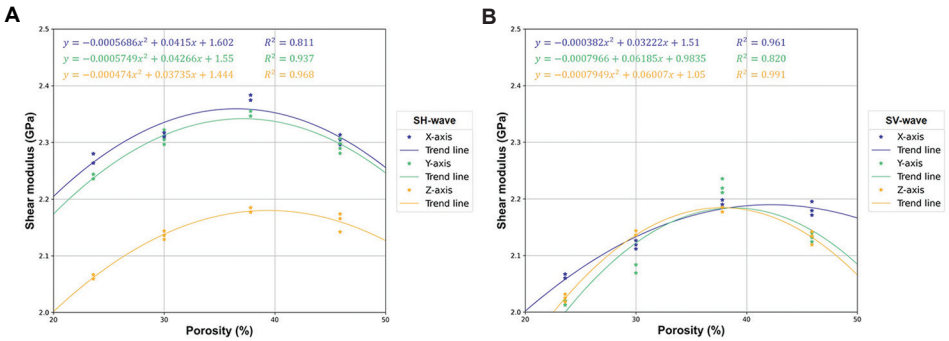



**Figure 13.** Shear modulus versus porosity in the permafrost simulation model calculated using (A) shear-horizontal (SH)-wave and (B) shear-vertical (SV)-wave velocities. The trend lines were generated using a second-order polynomial. Blue, green, and yellow markers represent the shear modulus measured along the X-, Y-, and Z-axes, respectively.

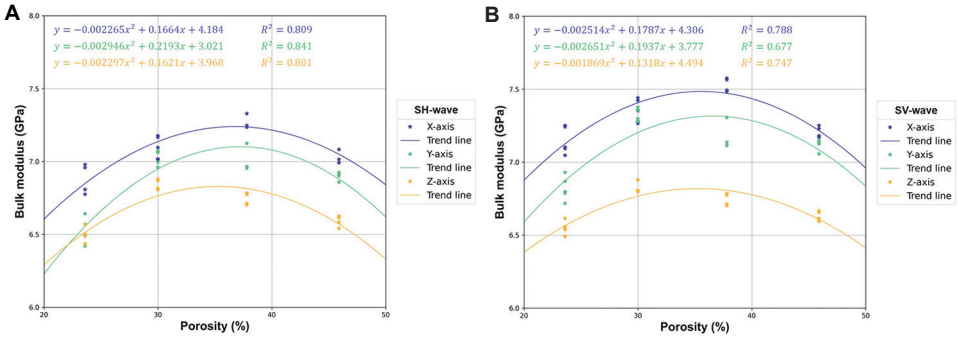



**Figure 14.** Bulk modulus versus porosity in the permafrost simulation model calculated using (A) shear-horizontal (SH)-wave and (B) shear-vertical (SV)-wave velocities. The trend lines were generated using a second-order polynomial. Blue, green, and yellow markers represent the shear modulus measured along the X-, Y-, and Z-axes, respectively.

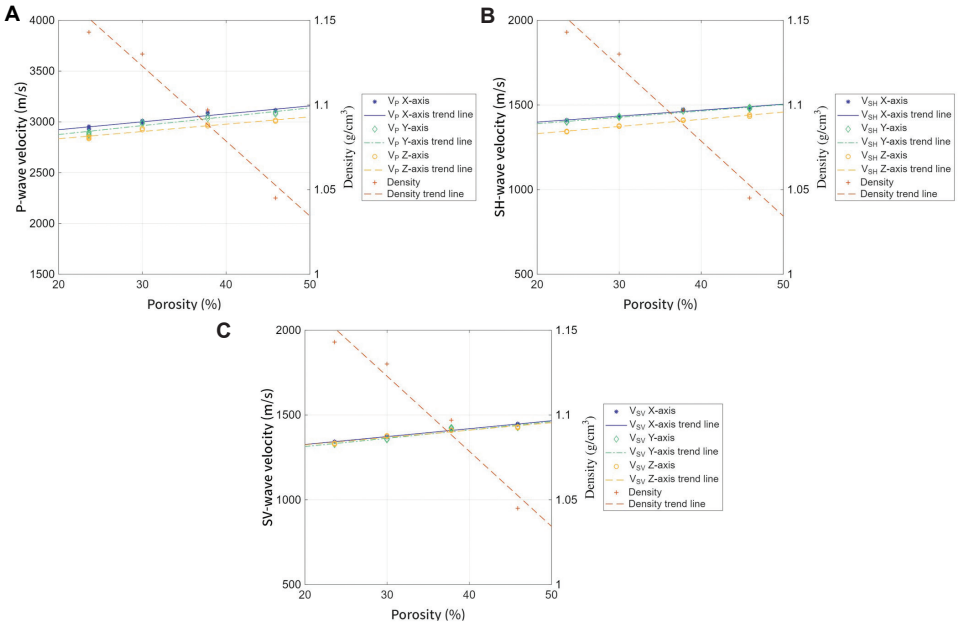



**Figure 15.** Relationships between seismic velocity/density and porosity in the permafrost simulation model. (A) P-wave velocity. (B) Shear-horizontal (SH)-wave velocity. (C) Shear-vertical (SV)-wave velocity. Blue, green, and yellow markers/lines represent the seismic wave velocities/trend lines along the X-, Y-, and Z-axes, respectively. The orange markers/lines represent the density/density trend lines.

velocity and density with respect to porosity were analyzed separately for each wave type. With increasing porosity, the seismic velocity increased due to the reinforcing effect of ice filling the larger pore volumes. In contrast, the overall density of the model decreased as porosity increased, attributable to the added pore volume reducing the bulk density of the material despite the presence of ice. Notably, the point of intersection between the seismic velocity and density trends corresponds to the porosity at which the elastic modulus begins to decline. This suggests that at low porosities, the increase in seismic velocity dominates the modulus calculation, increasing the elastic modulus. However, beyond a critical porosity, the reduction in density becomes the dominant factor, leading to a decrease in the elastic modulus despite the continued increase in velocity.

The effect of freezing was analyzed to understand its influence on the elastic behavior of the permafrost simulation model. When water within the pore spaces freezes, it undergoes a phase change accompanied by volume expansion. This freezing-induced expansion plays a complex role in determining the elastic modulus, as it can contribute to either strengthening or weakening of the material, depending on the porosity and structural response of the model.

In cases where the pore volume is limited, such as in low-porosity models, the expansion of ice during freezing

can generate compressive stress on the surrounding solid matrix. This stress enhances particle interlocking and improves the overall structural integrity of the material, thereby increasing the elastic modulus. In essence, the expansion serves to bind the structure more tightly by exerting pressure on the grain contacts, leading to increased stiffness and rigidity.

However, when porosity is relatively high, the extent of ice expansion becomes more significant, which can exceed the mechanical tolerance of the surrounding particles. In such cases, the excessive internal pressure may induce microcracks within the solid particles or at the grain boundaries.[33,34] These microcracks act as zones of mechanical weakness, which degrade the material's stiffness and, consequently, reduce the elastic modulus.

Figure 16 presents a schematic diagram illustrating this mechanism, where excessive volume expansion caused by freezing leads to the development of microcracks in the particle structure. This explanation aligns with the trend seen in the experimental results: when porosity was low, the elastic modulus increased with increasing porosity; when porosity increased beyond the threshold, the elastic modulus decreased with increasing porosity.

## 5. Discussion

This study demonstrated the feasibility of using FDM-based 3D printing technology to fabricate permafrost simulation models with controlled porosity, enabling a systematic investigation of the relationship between porosity and elastic modulus. By overcoming the limitations of using natural rock cores, the proposed method provided an effective experimental method for studying the elastic modulus of permafrost with respect to porosity.

The key finding of this study is that the elastic modulus does not increase monotonically with porosity under fully ice-saturated conditions. Instead, the modulus increases up to a critical point, after which it begins to decline. This non-linear behavior is primarily attributed to two competing effects: (i) the increase in seismic velocity due to enhanced ice bonding within pore spaces at lower porosity levels, and (ii) the decrease in bulk density and potential formation of microcracks induced by freezing-related volumetric expansion at higher porosity levels. The intersection point of the velocity and density trends, as shown in Figure 15, aligns closely with the porosity at which the elastic modulus reaches its peak, supporting this interpretation.

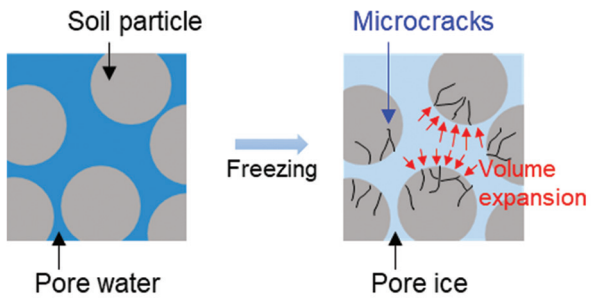



**Figure 16.** Schematic diagram of microcracks caused by ice volume expansion

These results highlight the dual role of freezing in permafrost systems. While ice formation can enhance the mechanical stiffness of the material by binding particles together, excessive ice expansion in high-porosity media may generate internal stresses that lead to microstructural damage, ultimately reducing the strength. This is quantitatively supported by the porosity measurements presented in Table 2. While the lower-porosity models (Models 1 and 2) exhibited negligible variations in porosity before and after freezing, the higher-porosity models (Models 3 and 4) showed more pronounced discrepancies. These deviations suggest that the internal stress generated by the volumetric expansion of ice caused structural damage in the high-porosity frameworks. Consequently, this induced damage is interpreted as the primary cause of the formation of microcracks and the subsequent decrease in elastic modulus observed beyond the critical porosity threshold.

The experimental method proposed here also opens the door for broader applications. By replacing ice with other fluids, such as water, oil, and brine, this technique can be used to study the porosity-dependent behavior of various saturated porous media. Moreover, varying the shape, size, and distribution of pores in models would allow for more comprehensive assessments of how pore architecture influences elastic properties.

Nevertheless, some limitations of the study should be acknowledged. The 3D-printed model, while effective for simulating elastic wave behavior, does not fully replicate the complex mineralogical and thermal properties of natural permafrost. In addition, freezing conditions in the experiment were static and uniform, whereas *in situ* freezing processes in natural environments are often dynamic and heterogeneous. Therefore, future research necessitates a comparative analysis between the results of laboratory experiments and those obtained from actual permafrost to bridge the gap between the simulation models and natural conditions. Furthermore, this study was conducted under conditions of complete ice saturation and zero salinity. However, natural permafrost often exhibits incomplete ice saturation and may contain saline pore water. Finally, although key properties such

as ice saturation, seismic velocity, and elastic moduli are sensitive to pressure and temperature changes, this study did not account for these environmental variables. Future research should incorporate more realistic boundary conditions, multi-phase modeling, and time-dependent effects to better simulate real-world permafrost behavior. In particular, further studies considering varying degrees of ice saturation, salinity, pressure, and temperature conditions are recommended to provide a more comprehensive analysis.

In summary, this study offers a novel experimental approach to permafrost research through the use of 3D printing and seismic physical modeling. It provides new insights into the mechanical response of frozen porous media and establishes a foundation for further research into the stability and geophysical characterization of permafrost environments.

## 6. Conclusion

This study proposed a method for controlling porosity and simulating permafrost using 3D printing. Furthermore, a permafrost simulation model with varying porosity was used in seismic physical modeling to analyze the elastic modulus according to porosity. The porosity of the permafrost simulation model was varied by controlling the number of pores and printed using an FDM 3D printer. The printed model was saturated with ice to simulate permafrost, and data were collected using seismic physical modeling. To analyze the elastic modulus according to porosity, seismic velocity was calculated from data obtained from seismic physical modeling. The results showed that the elastic modulus increased to approximately 38% of the porosity and then decreased, indicating the seismic wave velocity–density relationship and volumetric expansion due to freezing.

This study proposed a method utilizing 3D printing to overcome the limitations of controlling porosity in natural rock cores for permafrost research. Although the 3D-printed models differ from natural rocks, they demonstrate that the ice expansion within pores significantly affects the elastic modulus. From an engineering perspective, these findings provide a quantitative basis for evaluating the stability of infrastructure constructed on permafrost. By identifying the critical porosity threshold at which mechanical performance begins to degrade due to freezing expansion, engineers can better predict potential weakness zones and assess the bearing capacity of pipelines and foundations. Saturating the 3D-printed porous model, used as a permafrost simulation model, with fluids other than ice allows for the analysis of fluid properties associated with porosity. Varying pore structure and size enable systematic investigation into the effects of porosity within permafrost environments. Moreover, the applicability of this model further extends to studying permafrost degradation. By submerging the frozen model in water and controlling the temperature to induce gradual phase changes, the thawing process can be simulated. This approach allows for the monitoring of changes in elastic properties during the transition from frozen to thawed states, yielding valuable data for research on thawing environments. This approach may be extended to geological models, with potential applications in seismic exploration of permafrost regions.

## Acknowledgments

None.

## Funding

This research was supported by the Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries, Korea (RS-2023-00259633) and by the "Busan Regional Innovation System and Education (RISE)" Project supported by the Ministry of Education and Busan Metropolitan City (2025-RISE-02-002-012).

## Conflict of interest

The authors declare that they have no competing interests.

## Author contributions

*Conceptualization:* Daechul Kim, Sungryul Shin
*Formal analysis:* Yeonjin Choi
*Investigation:* Daechul Kim
*Methodology:* Wookeen Chung
*Visualization:* Daechul Kim
*Writing–original draft:* Daechul Kim
*Writing–review & editing:* Daechul Kim, Sungryul Shin

## Availability of data

All data analyzed have been presented in the paper.

## References

1. Van Everdingen RO, ed *Multi-Language Glossary of Permafrost and Related Ground-Ice Terms in Chinese, English, French, German, Icelandic, Italian, Norwegian, Polish, Romanian, Russian, Spanish, and Swedish*. Germany: International Permafrost Association, Terminology Working Group; 1998.
2. Obu J. How much of the earth's surface is underlain by permafrost? *J Geophys Res Earth Surf*. 2021;126(5):2021JF006123.

   doi: 10.1029/2021JF006123

3. Schuur EA, McGuire AD, Schädel C, *et al*. Climate change and the permafrost carbon feedback. *Nature*. 2015;520: 171-179.

doi: 10.1038/nature14338

4. Biskaborn BK, Smith SL, Noetzli J, *et al*. Permafrost is warming at a global scale. *Nat Commun*. 2019;10:264.

doi: 10.1038/s41467-018-08240-4

5. Hjort J, Streletskiy D, Doré G, Wu Q, Bjella K, Luoto M. Impacts of permafrost degradation on infrastructure. *Nat Rev Earth Environ*. 2022;3:24-38.

doi: 10.1038/s43017-021-00247-8

6. Zhang Z, Zhang J, Chen J, Luo J, Lei B. Study on the stability of excavation process of permafrost subgrade slope in Alpine region. *Sci Rep*. 2025;15:4612.

doi: 10.1038/s41598-025-85651-6

7. Timur A. Velocity of compressional waves in porous media at permafrost temperatures. *Geophysics*. 1968;33(4): 584-595.

doi: 10.1190/1.1439954

8. Zimmerman RW, King MS. The effect of the extent of freezing on seismic velocities in unconsolidated permafrost. *Geophysics*. 1968;51(6):1285-1290.

doi: 10.1190/1.1442181

9. Dou S, Nakagawa S, Dreger D, Ajo-Franklin J. A rock-physics investigation of unconsolidated saline permafrost: P-wave properties from laboratory ultrasonic measurements. *Geophysics*. 2016;81(1):WA233-WA245.

doi: 10.1190/geo2015-0176.1

10. Li B, Matsushima J. Influence of ice properties on wave propagation characteristics in partially frozen soils and rocks: A temperature-dependent rock-physics model. *Geophysics*. 2024;89(5):MR281-MR295.

doi: 10.1190/geo2023-0694.1

11. Kurfurst PJ. Ultrasonic wave measurements on frozen soils at permafrost temperatures. *Can J Earth Sci*. 1976;13(11): 1571-1576.

doi: 10.1139/e76-163

12. Zhang R, Ding P, Zhang F, Li J. Innovative 2D ultrasonic reflection measurements in partially frozen unconsolidated sediment. *Geophysics*. 2025;90(4):MR295-MR306.

doi: 10.1190/geo2024-0657.1

13. ASTM International. *ISO/ASTM 52900:2015 Additive Manufacturing - General Principles - Terminology*. Washington, D.C: ASTM International; 2015.

doi: 10.1520/ISO ASTM 52900-15

14. Ishutov S, Hasiuk FJ, Harding C, Gray JN. 3D printing sandstone porosity models. *Interpretation*. 2015;3(3):SX49-SX61.

doi: 10.1190/INT-2014-0266.1

15. Ibrahim ER, Jouini MS, Bouchaala F, Gomes J. Simulation and validation of porosity and permeability of synthetic and real rock models using three-dimensional printing and digital rock physics. *ACS Omega*. 2021;6(47):31775-31781.

doi: 10.1021/acsomega.1c04429

16. Ding P, Zhang F, Li X, Chai Y. Analyzing the seismic response characteristics of P, SV, and SH waves to reservoir parameters using physical modeling. *Geophysics*. 2024;89(5):M109-M121.

doi: 10.1190/geo2023-0454.1

17. Ding P, Zhang F, Chai Y, Luo P, Gao L. Seismic physical modeling study for a shale oil reservoir with complex lithofacies and meter-scale structures. *Geophysics*. 2023;88(4):B179-B193.

doi: 10.1190/geo2022-0582.1

18. Huang L, Dyaur N, Stewart RR. Elastic properties of 3D-Printed physical models: fluid substitution observations in cracked media. In: *SEG Technical Program Expanded Abstracts 2015. Society of Exploration Geophysicists Annual Meeting Technical Program 2015*; 2015. p. 3100-3104.

doi: 10.1190/segam2015-5915204.1

19. Huang L, Stewart RR, Dyaur N, Franceschi JB. 3D-printed rock models: Elastic properties and the effects of penny-shaped inclusions with fluid substitution. *Geophysics*. 2016;81(6):D669-D677.

doi: 10.1190/geo2015-0655.1

20. Dande S, Stewart RR, Dyaur N. The effect of fluids and their viscosity on the elastic-wave velocity and anisotropy of 3D-printed VTI rock models. In: SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists Annual Meeting Technical Program 2018; 2018. p. 347-350.

doi: 10.1190/segam2018-2998577.1

21. Dande S, Stewart RR, Dyaur N, Franceschi JB. The Effect of Proppant-Filled Fractures on the Elastic Properties of 3D-Printed Rock Models. In: SEG Technical Program Expanded Abstracts 2017. Society of Exploration Geophysicists Annual Meeting Technical Program; 2017. 2017. p. 5767-5771.

doi: 10.1190/segam2017-17789169.1

22. Wang J, Stewart RR, Dyaur NI. Seismic Response Analysis of a 3D-Printed Dual-Porosity Physical Model: Marine Case. In: SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists Annual Meeting Technical Program; 2018. p. 301-305.

doi: 10.1190/segam2018- 2997200.1

23. Zerhouni O, Tarantino MG, Danas K. Influence of the

Internal Geometry on the Elastic Properties of Materials using 3D-Printing of Computer-Generated Random Microstructures. In: SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists Annual Meeting Technical Program 2018; 2018. p. 3713-3718.

doi: 10.1190/segam2018-2998182.1

24. Kim D, Shin S, Chung W. Analysis of Seismic Characteristics for Permafrost According to Porosity Ratio using 3D Printing Technology. *Presented at*: *EGU General Assembly 2024*. Vienna, Austria: EGU24-7116; 2024.

doi: 10.5194/egusphere-egu24-7116

25. Pullan SE, MacAuley HA, Hunter JAM, Good RL, Gagne RM, Burns RA. Permafrost distribution determined from seismic refraction. In: Pelletier BR, ed *Marine Science Atlas of the Beaufort Sea - Geology and Geophysics*. Miscellaneous Report 40. Ottawa: Geological Survey of Canada; 1987. p. 37.

doi: 10.4095/126940

26. Redwood B, Schoffer F, Garret B. *The 3D Printing Handbook Technologies, Design and Applications*. Netherlands: 3D HUBS; 2017.

27. Zhao Y, Wang L, Li H. Effect of freeze-thaw on the stability of a cutting slope in a high-latitude and low-altitude permafrost region. *Appl Syst Innov*. 2020;3(3):36.

doi: 10.3390/asi3030036

28. Syas'ko V, Shikhov A. Assessing the state of structural foundations in permafrost regions by means of acoustic testing. *Appl Sci*. 2022;12(5):2364.

doi: 10.3390/app12052364

29. Wyllie MRJ, Gregory AR, Gardner LW. Elastic wave velocities in heterogeneous and porous media. *Geophysics*. 1956;21(1):41-70.

doi: 10.1190/1.1438217

30. Birch F. The velocity of compressional waves in rocks to 10 kilobars: 2. *J Geophys Res*. 1961;66(7):2199-2224.

doi: 10.1029/JZ066i007p02199

31. Wang YH, Lo KF, Yan WM, Dong XB. Measurement biases in the bender element test. *J Geotech Geoenviron Eng*. 2007;133(5):564-574.

doi: 10.1061/(ASCE)1090-0241(2007)133:5(564)

32. Han Q, Wang Z, Teng Z, Zhang Y, Cao J. Analysis on "three-box" model of stress-strain in frozen soil porous media based on representative macroscopic Volume. *Front Earth Sci*. 2022;10:991384.

doi: 10.3389/feart.2022.991384

33. Carvalho FCS, Chen CN, Labuz JF. Measurements of effective elastic modulus and microcrack density. *Int J Rock Mech Min Sci*. 1997;34(3-4):43.e1-43.e11.

doi: 10.1016/S1365-1609(97)00239-6

34. Wang T, Jia H, Sun Q, Lu T, Tang L, Shen Y. Pressure melting of pore ice in frozen rock under compression. *Cold Reg Sci Technol*. 2023;210:103768.

doi: 10.1016/j.coldregions.2023.103856

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Seismic characterization and spatial distribution analysis of coal gangue in the deep coalbed of the Northeastern Ordos Basin

**Zelei Jiang[1,2*], Xuri Huang[1,2,3,4], Dong Zhang[5], Yucong Huang[1,2], and Yong Wu[1,2]**

[1]State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, School of Geosciences and Technology, Southwest Petroleum University, Chengdu, Sichuan, China

[2]United Research Institute of Seismic Acquisition and Imaging Method, School of Geosciences and Technology, Southwest Petroleum University, Chengdu, Sichuan, China

[3]Natural Gas Geology Key Laboratory of Sichuan Province, School of Geosciences and Technology, Southwest Petroleum University, Chengdu, Sichuan, China

[4]Key Laboratory of Piedmont Zone Oil and Gas Geophysical Exploration Technology for Petroleum and Chemical Industry, School of Geosciences and Technology, Southwest Petroleum University, Chengdu, Sichuan, China

[5]Department of Geophysical, Exploration and Development Research Institute, PetroChina Changqing Oilfield Company, Xi'an, Shaanxi, China

***Corresponding author:**
Zelei Jiang
(jiangzl9602@163.com)

**Citation:** Jiang Z, Huang X, Zhang D, Huang, Wu Y. Seismic characterization and spatial distribution analysis of coal gangue in the deep coalbed of the Northeastern Ordos Basin. *J Seismic Explor*. 2026;35(2):025490123. doi: 10.36922/JSE025490123

**Received:** December 4, 2025

**Revised:** January 21, 2026

**Accepted:** January 22, 2026

**Published online:** March 6, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Abstract

Deep coalbed methane exploration in the Ordos Basin holds significant potential, yet the frequent development of coal gangue severely hinders productivity and complicates horizontal well deployment. Accurately characterizing the spatial distribution of these thin coal gangue layers remains a critical challenge, particularly when constrained by limited seismic resolution and sparse well control. To address this challenge, we propose a novel paleogeomorphology-constrained stochastic seismic inversion workflow applied to the Benxi Formation in the YL area. This approach integrates petrophysical analysis with paleogeomorphic restoration, identifying natural gamma as the most sensitive lithological indicator. Under hierarchical geomorphic constraints, geostatistical simulation was utilized to predict the three-dimensional spatial probability of coal gangue occurrence. Quantitative validation demonstrates a robust correlation between coal gangue thickness and natural gamma response, with the inversion results achieving an accuracy of 81.82% in blind well validation. Spatially, gangue development is controlled by paleotopography, with higher probabilities concentrated in paleo-highs and slopes associated with stronger hydrodynamic conditions, while paleo-depressions exhibit superior coal continuity. This study not only overcomes the resolution limitations of traditional inversion in sparse-well areas but also provides a rigorous quantitative geological basis for sweet-spot identification and trajectory optimization in deep coalbed methane development.

***Keywords:*** Ordos Basin; Coal gangue-sensitive log parameter; Seismic characterization of deep coal gangue; Coal gangue distribution pattern; Deep coalbed methane

## 1. Introduction

In the Ordos Basin, coal-bearing strata were extensively developed in the Late Paleozoic transitional marine–continental environment. The main coalbeds, including the No. 5 coal of the Shanxi Formation (Lower Permian) and the No. 8 coal of the Benxi Formation (Upper Carboniferous), are rich in gas resources and represent the primary targets for deep coalbed methane (CBM) exploration. Unlike conventional CBM, coal-rock gas exhibits geological characteristics of high pressure, temperature, gas content, saturation, and free gas content.[1] Owing to differences in coal-forming periods and depositional environments, coal gangue layers commonly occur within deep coalbeds, which not only affect gas content but also complicate the design of horizontal well trajectories.[2]

Previous studies have investigated coal gangue identification and characterization mainly based on log data, sedimentological observations, and laboratory experiments. Guo *et al.*[3] reported that coal gangue generally exhibits higher density and natural gamma (GR) responses than coalbed, reflecting their mudstone-dominated or siltstone-dominated compositions. Zhang *et al.*[4] further demonstrated that coal gangue appears as distinct dark bands on image logs, characterized by elevated thorium and uranium values, higher density, and lower resistivity compared with pure coalbeds. These logging indicators provide a foundation for distinguishing coal gangue from coal in vertical profiles.

From a reservoir perspective, Sun *et al.*[5] found that the spatial variability and complexity of coal gangue development strongly influence reservoir properties and gas-bearing capacity. Specifically, coal gangue increases ash yield, reduces methane adsorption capacity, disrupts coal continuity, and complicates coalbed structures, ultimately hindering CBM preservation and production. Considerable efforts have also been devoted to examining the mechanical effects of coal gangue layers and surrounding rocks. Li *et al.*[6] conducted numerical simulations and revealed that coal gangue thickness directly affects stress concentration within coal-bearing strata, influencing bed stability. Yan *et al.*[7] conducted rock mechanical experiments and stress-field modeling and reported that increasing coal gangue thickness reduces both the maximum horizontal stress and the overall stress difference, thereby altering hydraulic fracturing behavior and fracture propagation patterns. Huang *et al.*[8] analyzed stress distribution and failure characteristics in coal gangue composite strata and found that coal gangue tends to bear relatively higher horizontal stress owing to its greater mechanical strength.

With technological advances, computed tomography scanning has also been introduced for coal gangue identification. Gong *et al.*[9,10] systematically investigated the pressure-dependent elastic anisotropy of deep coal reservoirs by integrating microscopic observations, ultrasonic experiments, and anisotropic rock physics modeling, thereby providing reliable support for stress simulation applications in deep coal reservoirs. Wu *et al.*[11] conducted gas desorption experiments combined with multiscale pore structure analyses and found that soft coal exhibited more developed pore structures and better pore connectivity, thereby demonstrating that high-rank soft coal has a significantly higher initial gas desorption capacity than hard coal. Wu *et al.*[12] proposed a rapid method for determining residual gas pressure in pre-pumped coal seams based on adsorption–desorption equilibrium experiments, offering higher efficiency, fewer required parameters, and lower errors than conventional Langmuir-based indirect methods and direct downhole pressure measurements.

In addition, Zhang *et al.*[13] found that cyclic loading–unloading stress can significantly enhance the permeability of deep coalbed gas reservoirs, with the most pronounced effect observed in coal seams containing inclined fractures, providing an important mechanical basis for optimizing the extraction of complex coalbed gas. In another study, Huang *et al.*[14] used spiral computed tomography to achieve high-resolution detection of coal gangue layers and reported that coalbeds with lower coal gangue content and thinner single layers exhibit higher gas content and greater development potential. Yang *et al.*[15] proposed an attribute-fusion method integrating vibration signals with cross-optimal selection to improve coal gangue identification accuracy. In the Ordos Basin, Xu *et al.*[16] analyzed logs from the No. 8 coal and classified the coalbed into three structural types—single-split, double-split, and triple-split—providing essential insights into internal heterogeneity.

Although existing studies have characterized coal gangue properties from various perspectives, research integrating paleogeomorphology, log, and seismic inversion to reveal its three-dimensional (3D) spatial distribution remains limited, highlighting the necessity of the present work.

Overall, existing studies on coal gangue have primarily focused on log-based identification, pore structure characterization, and mechanical properties, whereas insufficient attention has been paid to the fine-scale spatial distribution of coal gangue at inter-well and regional scales. In particular, under a 3D geological framework, the continuity, heterogeneity, and spatial coupling between coal gangue distribution, paleogeomorphology, and depositional microfacies remain poorly constrained,

limiting a comprehensive understanding of coal gangue development patterns and their implications for CBM accumulation and development. In practice, the coal gangue directly affects well trajectory design and perforation strategies for reservoir production.[17]

In this study, log and core data were employed to achieve high-precision classification of coal and gangues, with GR identified as the most sensitive parameter for coal gangue recognition. Seismic characterization was further applied to delineate the spatial distribution of coal gangue within the No. 8 coalbed. Combined with regional tectonic evolution and sedimentary facies analysis, the depositional features and distribution patterns of coal gangue were clarified, providing significant guidance for sweet-spot identification and horizontal well trajectory optimization in deep CBM development within the Ordos Basin.

## 2. Geological overview of the study area

The YL area (Yulin area) is situated in the northeastern part of the Yishan Slope within the Ordos Basin (Figure 1). The top boundary of the No. 8 coalbed in the Benxi Formation presents a gentle monocline structure dipping from east to west, reflecting the overall tectonic framework of the region. The burial depth of the coalbed gradually increases westward, ranging from 2,260 m to 3,250 m,[18] indicating progressive subsidence during the late Paleozoic. The No. 8 coalbed was deposited in a lagoon–tidal flat environment,[19] where relatively stable hydrodynamic conditions favored the accumulation and preservation of thick coal seams. This coalbed is widely distributed across the basin and hosts abundant CBM resources, making it a key target for exploration and development. Its thickness generally increases toward the east and decreases toward the southwest. In the central and eastern portions of the study area, the coalbed thickness ranges from 3 m to 12 m, with an average of 7.2 m.

Due to fluctuations in relative sea level, multiple short-term depositional cycles were developed within the Benxi Formation.[20] These cycles caused repeated transitions between coal-forming environments and mudflat sedimentation. Consequently, thin mudstone layers were frequently identified as coal gangue within the No. 8 coalbed, forming numerous coal gangue layers of varying thickness. These coal gangue layers display significant lateral and vertical variability, both between wells and within individual well profiles. Such variability results in rapid spatial changes in coalbed architecture, which significantly enhances reservoir heterogeneity.[21] The distribution and continuity of these coal gangue layers not only affect the mechanical properties of the coalbed but also exert a substantial influence on its internal permeability structure and CBM enrichment patterns. Understanding their spatial characteristics is therefore crucial for accurately predicting reservoir

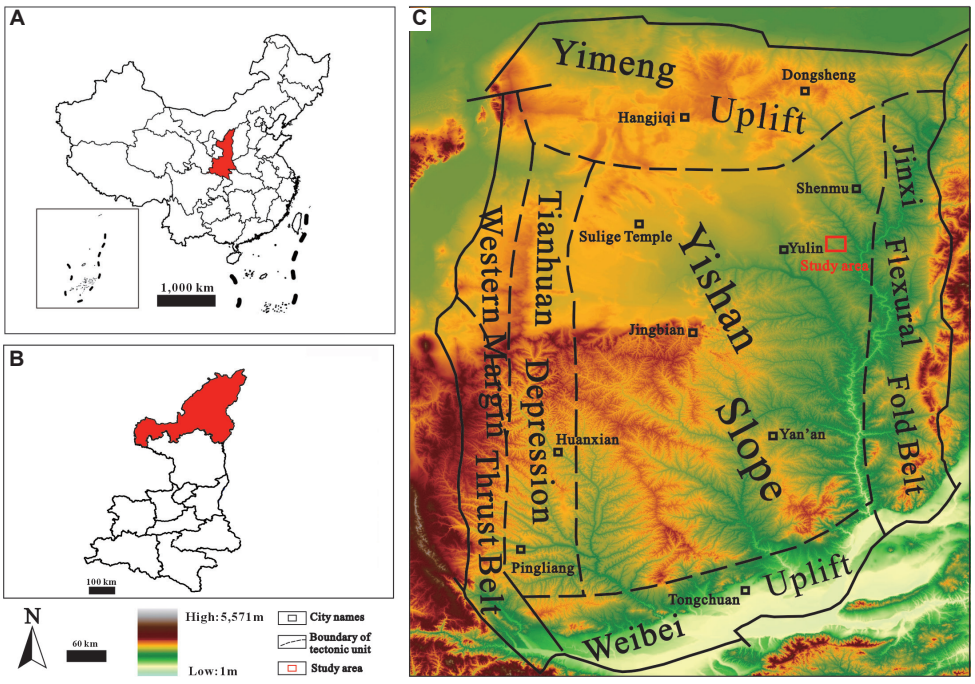



**Figure 1.** Location map of the study area. (A) Location of Shaanxi Province in China. (B) Location of the YL work area in Shaanxi Province. (C) Location of the YL work area in the elevation map of the Ordos Basin.

quality and optimizing CBM development strategies in the YL area.

The target zone of this study was the Carboniferous Benxi Formation, deposited in a characteristic marine–continental transitional environment. Drilling data and comprehensive log interpretations indicate that its lithology is dominated by sandstone, coal, dark mudstone, and locally developed limestone. To better understand the internal architecture of the coal-bearing strata, the vertical lithologic assemblages within the target zone were systematically analyzed based on regional sedimentary evolution patterns combined with detailed well and log responses. As shown in Figure 2, the Benxi Formation mainly represents a tidal-flat subfacies system, characterized by frequent shifts in hydrodynamic conditions and repeated alternations of clastic and organic-rich deposits.

The coalbeds within the zone display distinctive log signatures, including low GR, density, and acoustic velocity, which collectively reflect their high organic content and relatively weak compaction. In contrast, limestone exhibits the highest velocity and density among all lithologies due to its strong cementation and carbonate composition. Sandstone and mudstone share relatively similar density and velocity ranges. However, they can be effectively differentiated using GR values, with sandstone typically showing significantly lower gamma responses. Coal gangue, primarily composed of thin interbedded mudstone, exhibits higher acoustic velocity and density than coalbeds and demonstrates logging characteristics similar to mudstone, enabling their identification through sensitive parameter comparison.

With the deepening exploration and large-scale development of deep CBM resources in the Ordos Basin, there is an increasing need for refined characterization of the internal structure, continuity, and spatial heterogeneity of coalbeds. In particular, the frequent occurrence of coal gangue layers poses significant challenges for reservoir evaluation, horizontal well drilling, and production performance prediction. Therefore, a systematic investigation into the spatial distribution and geological controls of coal gangue has become essential. Through the integrated analysis of sensitive log parameters, combined with seismic characterization, sedimentary evolution research, and facies interpretation, this study aims to elucidate the distribution patterns and controlling mechanisms of coal gangue within the Benxi Formation. The findings will provide a robust geological foundation and technical support for high-precision sweet-spot identification, optimized horizontal well placement, and the efficient development of deep CBM resources in the Ordos Basin.

## 3. Seismic characterization of deep coal gangue constrained by log patterns and paleogeomorphological boundaries

### 3.1. Analysis of log-sensitive parameters for coal gangue

Based on log data from six representative wells in the study area, combined with drilling and core data, detailed logging response analyses were conducted for both coalbed- and coal gangue-bearing intervals (Figure 3). Eight representative logging parameters were selected to systematically compare the petrophysical differences between coalbed and coal gangue layers and identify diagnostic response patterns that can reliably characterize coal gangue development.

The comparative analysis revealed that coal gangue layers generally exhibited a typical "five-high and three-low" log response pattern. Specifically, coal gangue showed high corrected gamma ray, density, acoustic velocity,

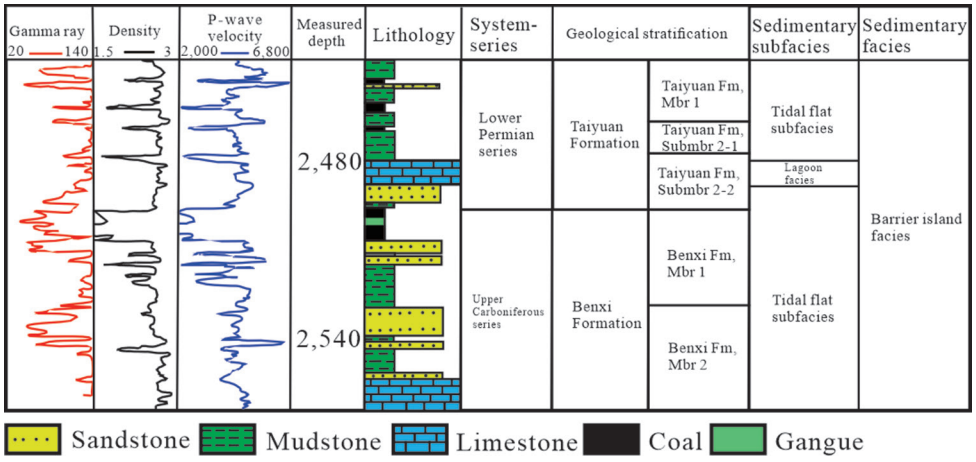



**Figure 2.** Representative well logging curves and the corresponding sedimentary facies map

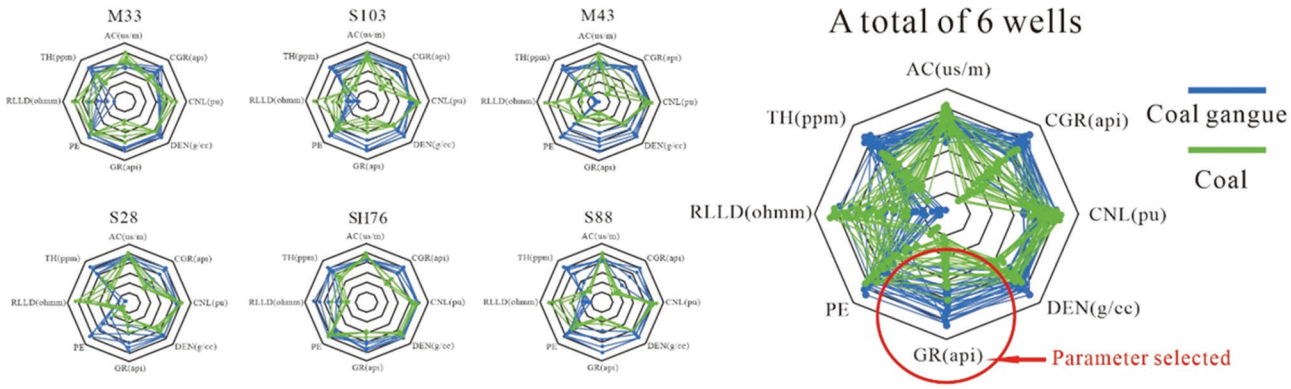



**Figure 3.** Comparison of log data-sensitive parameters between coalbeds and coal gangue. Radar charts for wells M33, S103, M43, S28, SH76, and S88 show sensitive log parameter responses for coal gangue layers in individual wells. The combined radar chart of all six wells presents the superimposed sensitive log parameter responses for coal gangue layers across all wells.
Abbreviations: AC: Acoustic transit time; CGR: Corrected gamma ray; CNL: Neutron porosity; DEN: Density; GR: Natural gamma; PE: Photoelectric absorption index; RLLD: Deep resistivity; TH: Thorium.

GR, and thorium content, accompanied by low deep resistivity, photoelectric absorption index, and neutron porosity. These characteristics reflect the higher clay and mud content of coal gangue, as well as their differences in mineral composition and pore structure relative to coal.

Among all parameters, the GR log demonstrated the highest sensitivity to coal gangue development. GR not only clearly differentiated mudstone coal gangue from coalbed but also provided stable and continuous responses across different wells. Compared with other indicators, GR showed the strongest correlation with the lithological variations observed in core samples, making it the most reliable parameter for delineating the thickness, boundaries, and spatial occurrence patterns of coal gangue layers. Therefore, GR was selected as the key sensitive attribute for subsequent seismic stochastic inversion aimed at generating the 3D probability volume of coal gangue distribution. This parameter choice ensures that the inversion results adequately reflect the vertical and lateral heterogeneity of coal gangue within the No. 8 coalbed, providing a solid foundation for subsequent reservoir characterization and sweet-spot prediction.

### 3.2. Probabilistic inversion of coal gangue spatial distribution

A 3D geological model of coalbeds and coal gangue provides a quantitative framework for characterizing their geological features and spatial distribution. It serves as an effective approach for fine-scale characterization of coal gangue and is also critical for optimizing horizontal well deployment and enhancing CBM development efficiency. By quantitatively characterizing coal gangue distribution in 3D space, such a model offers a more objective description of spatial patterns, thereby providing a robust basis for subsequent studies. Figure 4 shows the workflow of seismic characterization for coal gangue identification.

In geostatistical stochastic simulation, the two most critical tasks are: (i) Summarizing the statistical distribution of regionalized variables and (ii) Conducting variogram analysis. The variogram, as the fundamental tool of geostatistics, describes spatial correlation based on the principle that values at closer locations are more strongly correlated than those farther apart. The separation distance between two points is referred to as the lag ($h$), and the relationship between the variogram and lag distance $h$ is illustrated in the variogram plot. Key parameters include the range, nugget, and sill (Figure 5). Common theoretical variogram models applied in geological modeling include the exponential, spherical, and Gaussian models, all of which asymptotically approach the sill.[22]

Among them, the Gaussian model gradually approaches the sill ($c$), reaching $h$ = 0.95c in the range ($a$). This model is parabolic near the origin, making it the only variogram model with an inflection point. Its slope at the origin is zero, indicating very slow changes at short lag distances followed by rapid increases as the lag distance grows. The Gaussian model produces smoother and more continuous results at short ranges, but its overall stability is relatively lower compared with that of other models (**Equation 1**).

$$y(h) = c\left[1 - exp\left(-\frac{3h^2}{a^2}\right)\right] \quad (1)$$

Random simulation algorithms—constrained by input data—generate multiple geological realizations through stochastic functions while simultaneously considering

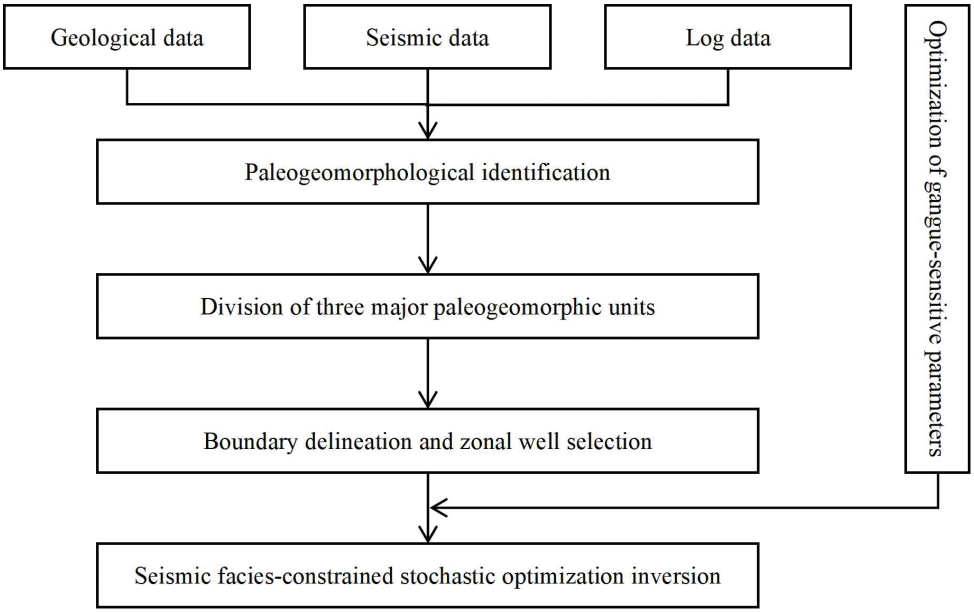



**Figure 4.** Workflow of seismic characterization of coal gangue

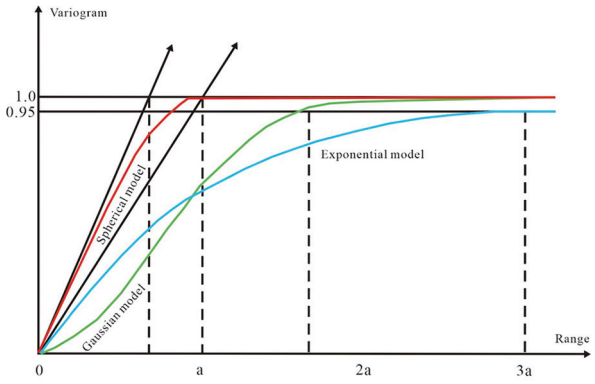



**Figure 5.** Common theoretical variogram models illustrating the relationship between variogram value and range

variable randomness and 3D space correlation. Sequential Gaussian simulation, a Kriging-based conditional simulation method for continuous variables, was applied in this study. The approach assumes that the input variables follow a Gaussian distribution and performs sequential simulation using Gaussian probability theory.

Before simulation, input data were detrended to conform to a standard normal distribution, and the final realizations were back-transformed to ensure consistency with the original data. In sequential Gaussian simulation, the study area is discretized into grids, and each grid node is simulated sequentially. At each node, the simulated value is determined by the mean and variance derived from the Kriging system and expressed as a cumulative conditional probability. Each realization of the regionalized variable $Z(x)$ is regarded as a random realization of the composite Gaussian distribution function $F(x)$, conditioned on $n$ known data points $Z(x_i)$ ($i$ = 1, 2, ..., $n$) and $m-1$ previously simulated values $Z(x_j)$ ($j$ = 0, 1, ..., $m-1$). In subsequent work, sequential indicator simulation will be integrated with stochastic inversion results to construct the coal gangue model.[23]

In this study, well logs, 3D seismic, and geological data from the YL area were integrated. For the No. 8 coalbed of the Benxi Formation, GR was selected as the coal gangue-sensitive inversion parameter. The distribution pattern of GR and its correlation with coal gangue development were analyzed based on log data and geological data, from which a low-frequency trend and spatial structural model were established. Well constraints were introduced as prior conditions. A geostatistical stochastic optimization seismic inversion method was then applied, combining geostatistical simulation with seismic fitting.[24] This method primarily considers the probability density function of well log data.[25] Iterative updates were performed, preferentially re-simulating areas with larger mismatches while retaining well-fitted regions. This global–local alternating optimization ensures that the inversion simultaneously honors regional distributions and local structural variations.

After multiple iterations, a suite of equiprobable GR-based probability models for coal gangue distribution was generated (Figure 6). These models satisfy the following conditions: (i) Consistency with well logs and geostatistical characteristics; (ii) Amplitude and phase fitting of seismic data; and (iii) Coherence between global distribution

patterns and local structures. The inversion results provided a robust foundation for quantitatively assessing the spatial distribution of coal gangue and analyzing the depositional evolution of the Benxi Formation coalbed in the YL area.

Based on seismic profile interpretation and stratigraphic horizon picking, the paleogeomorphic relief during the deposition of the No. 8 coalbed in the Benxi Formation was reconstructed. Through detailed structural analysis and seismic attribute extraction, the paleogeomorphic characteristics of the YL area were accurately identified. The resulting paleogeomorphic framework revealed distinct topographic differentiation, allowing the study area to be divided into three first-order geomorphic units with clearly defined boundaries. Figure 7A presents the 3D spatial display map, whereas Figure 7B shows the planar display map, displaying the following geomorphic units:

(i) Paleo-uplift area: Located in the western part of the study area, this zone is characterized by relatively high paleotopographic elevations, representing regions favorable for concentrated hydrodynamic energy and sediment bypass

(ii) Paleo-slope area: Situated between the uplift and the depression, this transitional zone displays a gentle east–west dip and serves as the main pathway for sediment transport from highlands to low-lying regions

(iii) Paleo-depression area: Distributed across the eastern part of the study area. This low-relief region exhibits lower elevations and provides favorable conditions for fine-grained sediment accumulation and coal-forming peat development.

The reconstructed paleogeomorphic pattern offers essential geological constraints for understanding coal gangue distribution and for guiding deep CBM exploration.

The thickness of the target coal-bearing strata is used as an important geological constraint for paleogeomorphic reconstruction. Thickness variations are interpreted to primarily reflect differential accommodation space controlled by paleotopography under relatively stable sediment supply conditions. In the study area, systematic lateral variations in coal-bearing strata thickness reveal clear paleogeomorphic differentiation, with thicker successions corresponding to paleo-depressions and thinner successions associated with paleo-uplift or slope settings.

Based on Table 1, the paleogeomorphic configuration during the coal deposition period exerted a significant control on the generation and spatial distribution of coal gangue. Paleo-depression areas exhibited the lowest coal gangue frequency and were characterized by relatively continuous, intact coalbeds, making them favorable zones for subsequent drilling and development.

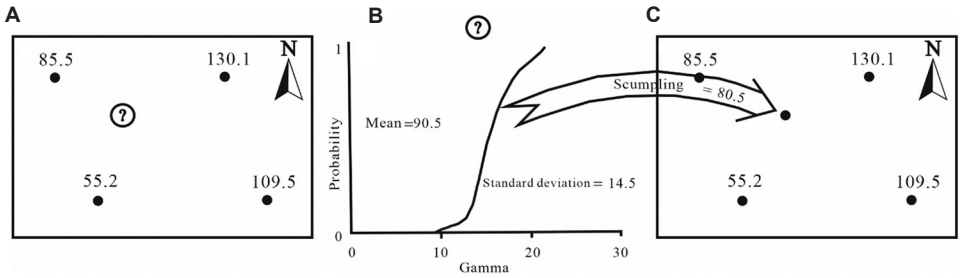



**Figure 6.** Schematic diagram of the stochastic simulation process. (A) Conditional data and locations to be simulated. (B) Grid node location. (C) Simulated value.

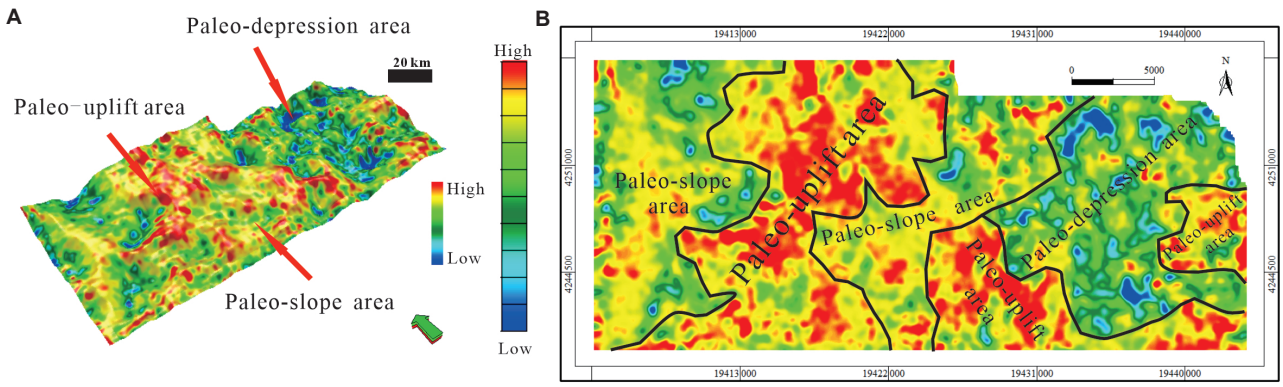



**Figure 7.** Paleogeomorphic region classification: (A) Three-dimensional visualization and (B) plan view

Conventional seismic inversion methods are often limited in complex depositional environments due to insufficient low-frequency information, and inversion constrained only by single wells suffers from uncertainties in low-frequency trend simulation, thereby affecting the accuracy of geological interpretation. To improve inversion precision, paleogeomorphic unit boundaries were introduced as constraints, serving as a crucial link between well data and depositional settings. By applying zonal constraints based on the three paleogeomorphic units, multi-scale control of seismic inversion was achieved.

In practice, inversion was performed using paleogeomorphic unit boundaries as geological control units, integrating well and seismic data for zoned inversion. Independent inversion calculations were performed for wells within different unit backgrounds. This approach preserved the lateral resolution of facies variations while effectively avoiding the "averaging effect" caused by excessively large vertical time windows, thereby significantly enhancing the fine-scale prediction capability of coal gangue distribution.

It should be noted that the zonal inversion applied to wells under different unit backgrounds refers to establishing inversion constraints separately within different paleogeomorphic zones, rather than spatially isolating the inversion calculations. All zones shared a unified inversion parameter system and the same initial model. This strategy ensures that the inversion results macroscopically adhere to regional geological depositional patterns while microscopically capturing local structural variations and heterogeneity, thereby significantly enhancing the spatial accuracy of coal gangue prediction in complex geological settings.

### 3.3. Blind-well validation using log data

Figure 8A illustrates the interpreted spatial distribution of coal gangue within the No. 8 coalbed based on log and well data, representing the actual geological configuration revealed by subsurface data. In comparison, Figure 8B displays the coal gangue probability map generated by the stochastic inversion results. Overall, the two maps show strong agreement in terms of spatial distribution patterns, lateral continuity, and the extent and location of major coal gangue-enriched zones. Areas with higher predicted coal gangue probability correspond closely to intervals confirmed by drilling to contain frequent coal gangue. This strong consistency validates the reliability and accuracy of the inversion method and indicates that the paleogeomorphically constrained inversion effectively captures the internal heterogeneity of the coalbed.

**Table 1. Number of coal gangue layers in wells within the three paleogeomorphic areas**

| Paleogeomorphic area | Number of coal gangue layers (well-based statistics) | Proportion of total wells in the area (%) |
|---|---|---|
| Paleo-depression area | 14 | 12.6 |
| Paleo-slope area | 33 | 42.3 |
| Paleo-uplift area | 52 | 64.9 |

Compared with multi-well interpolation results, which indicate consistent trends and more developed coal gangue in paleogeomorphic high areas, the lithofacies prediction based on 11 validation wells achieved an accuracy of 81.82% (Table 2). The discrepancies were mainly attributed to the extremely thin coal gangue layers, which may lead to misclassification.

To further validate the reliability of the inversion, a comparative analysis was conducted using seismic profiles and well-to-well correlation sections. Typical seismic sections of inversion parameters were compared with corresponding interpreted sections (Figure 9), and multi-well geological correlation profiles constructed were examined against inversion results (Figure 10). The validation results confirm that the predicted coal gangue locations, thicknesses, and trends closely match log-derived observations. This demonstrates the applicability of the stochastic optimization inversion method in the study area and confirms its accuracy in predicting coal gangue spatial distribution, thereby providing strong technical support for subsequent geological and reservoir analyses.

It is important to note that sparse well spacing has a significant impact on the lateral uncertainty of the prediction results. Since stochastic simulation relies on variograms to characterize spatial correlation, prediction uncertainty typically increases linearly with distance from well control points. In regions with relatively stable depositional environments, such as paleo-depressions, the error introduced by sparse well control is relatively minor. However, in highly heterogeneous transition zones, such as paleo-slopes, excessive inter-well spacing may limit the model's ability to capture high-frequency lithological variations. Therefore, the current prediction results (81.82% accuracy) represent the optimal estimation under existing well constraints, and data from future infill drilling are expected to further reduce lateral prediction uncertainty.

To address the limitation that stochastic inversion struggles to resolve thin layers due to seismic tuning thickness, future work will move beyond traditional statistical methods and focus on multi-attribute seismic pattern recognition. We plan to introduce machine learning algorithms, such as random forest or convolutional neural networks, to extract frequency-dependent attributes sensitive to thin layers and texture attributes sensitive to lithological boundaries. In addition, we aim to enhance the identification capability for

the "weak gangue signatures" described in the present study through non-linear mapping in high-dimensional feature space, thereby achieving a more refined characterization of subtle geological bodies under seismic resolution constraints.

## 4. Discussion

The Ordos Basin is a multi-cycle superimposed petroliferous basin with a basement of Archean–Paleoproterozoic metamorphic rocks. Structurally, it is an asymmetric north–south-trending rectangular basin, characterized by a steep and narrow western margin and a broad, gentle eastern margin. As illustrated in Figure 11, the basin has undergone five stages of tectonic and sedimentary evolution. The first stage occurred during Mesoproterozoic–Neoproterozoic rifting, dominated by large-scale intracontinental and marginal crustal disintegration. The second stage was the Early Paleozoic marginal marine basin stage, with relatively stable vertical movements and marine sedimentary deposition. The third stage was the Late Carboniferous–Middle Triassic intracratonic basin stage, during which the basin evolved from a cratonic epicontinental sea to an intracratonic depression, with sedimentation transitioning

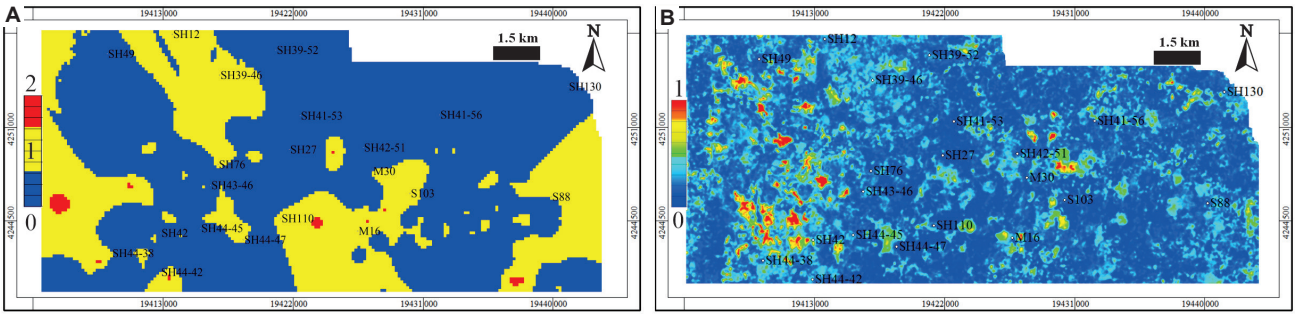



**Figure 8**. Spatial distribution of coal gangue. (A) Observed coal gangue based on log data and (B) predicted coal gangue probability distribution.

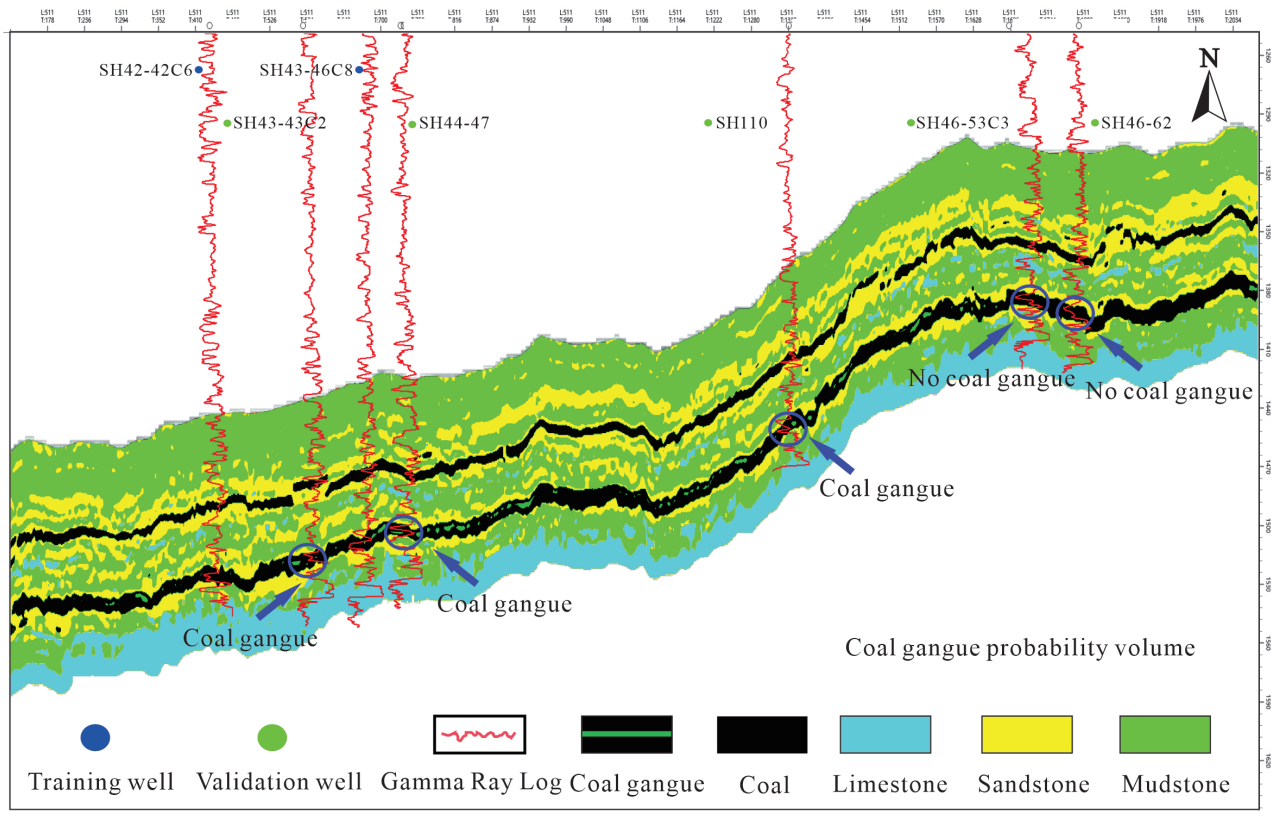



**Figure 9.** Seismic blind well profile verification

from marine–terrestrial transitional facies to lacustrine deposits. The fourth stage, spanning the Late Triassic and Early Cretaceous, was characterized by residual cratonic basin development, intense peripheral tectonism, thrust–fold deformation, tilting uplift, and regional unconformities between Triassic and Jurassic strata. During the Middle–Late Triassic extension, lacustrine systems reached their maximum development, depositing thick, high-quality source rocks. The fifth stage was the peripheral fault–depression stage, in which the basin shifted from subsidence to uplift and was subsequently segmented by faulting.[26,27]

**Table 2. Comparison of observed and predicted coal gangue in wells and match status**

| Validation well | Observed coal gangue | Predicted coal gangue | Match status | Accuracy |
|---|---|---|---|---|
| SH43-43C2 | Yes | Yes | Match | 81.82% |
| SH42 | Yes | No | Mismatch | |
| SH49 | Yes | Yes | Match | |
| SH44-38 | Yes | No | Mismatch | |
| SH44-41 | Yes | Yes | Match | |
| SH46-62 | No | No | Match | |
| SH44-47 | Yes | Yes | Match | |
| SH46-53C3 | No | No | Match | |
| M16 | Yes | Yes | Match | |
| SH110 | Yes | Yes | Match | |
| M30 | Yes | Yes | Match | |

Despite experiencing multiple tectonic movements throughout geological history, the Ordos Basin has remained structurally stable overall, characterized by the absence of large-scale active faults and relatively gentle deformation. Based on its current structural framework, the basin can be subdivided into six primary tectonic units: (i) The Yimeng Uplift in the north, (ii) The Weibei Uplift in the south, (iii) The Yishan Slope in the central part, (iv) The Jinxi Flexural Fold Belt to the east, (v) The Tianhuan Depression to the west, and (vi) The western thrust belt.

Among these, the Yishan Slope is the most extensive and significant unit, exerting a fundamental control on hydrocarbon enrichment patterns across the basin.[28] The Yishan Slope, developed during the late Early Cretaceous, forms a broad, gently west-dipping monocline extending approximately 400 km from north to south and 250 km from east to west. Its dip angle is typically <1°, although local structural variations, such as nose-like uplifts and minor warping, occur due to subtle tectonic adjustments.

The combination of wide areal extent, mild structural inclination, and high stability has made the Yishan Slope the principal zone for oil, gas, and CBM accumulation in the Ordos Basin. Under this gentle and stable tectonic background, sedimentation during the Early Carboniferous to Permian periods was dominated by epeiric sea environments. The initial transgression facilitated the deposition of subtidal carbonates across extensive shallow-marine platforms. As relative sea level gradually regressed, depositional

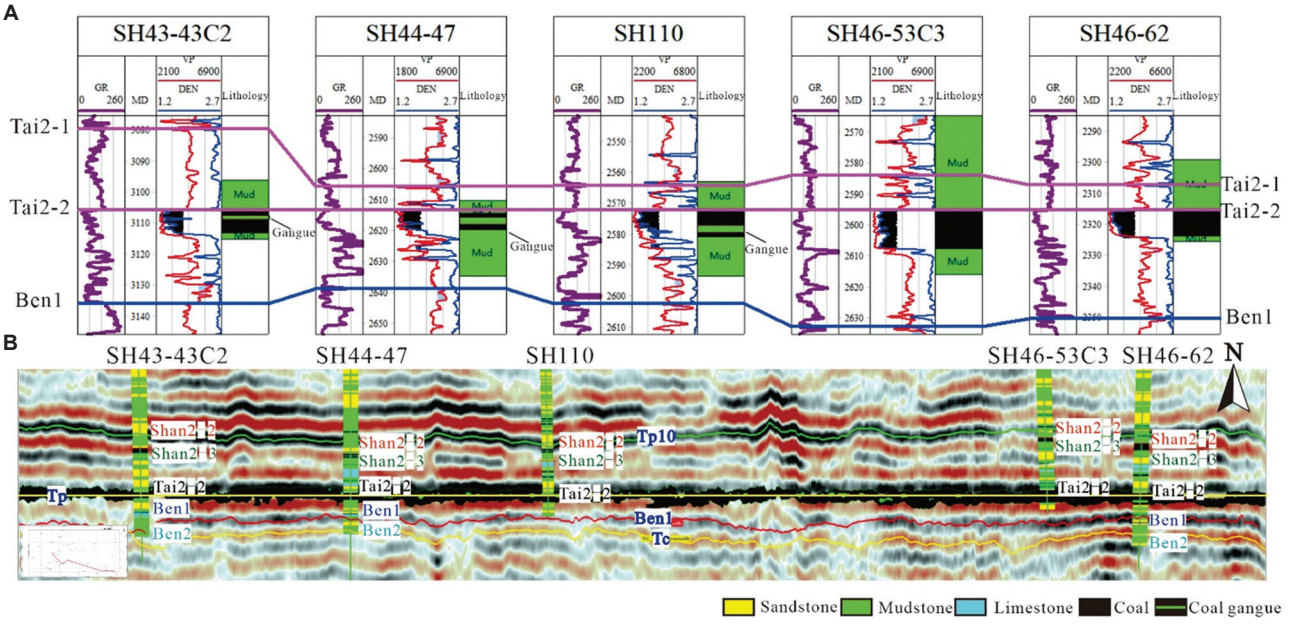



**Figure 10.** (A and B) Geological blind well cross-well profile verification

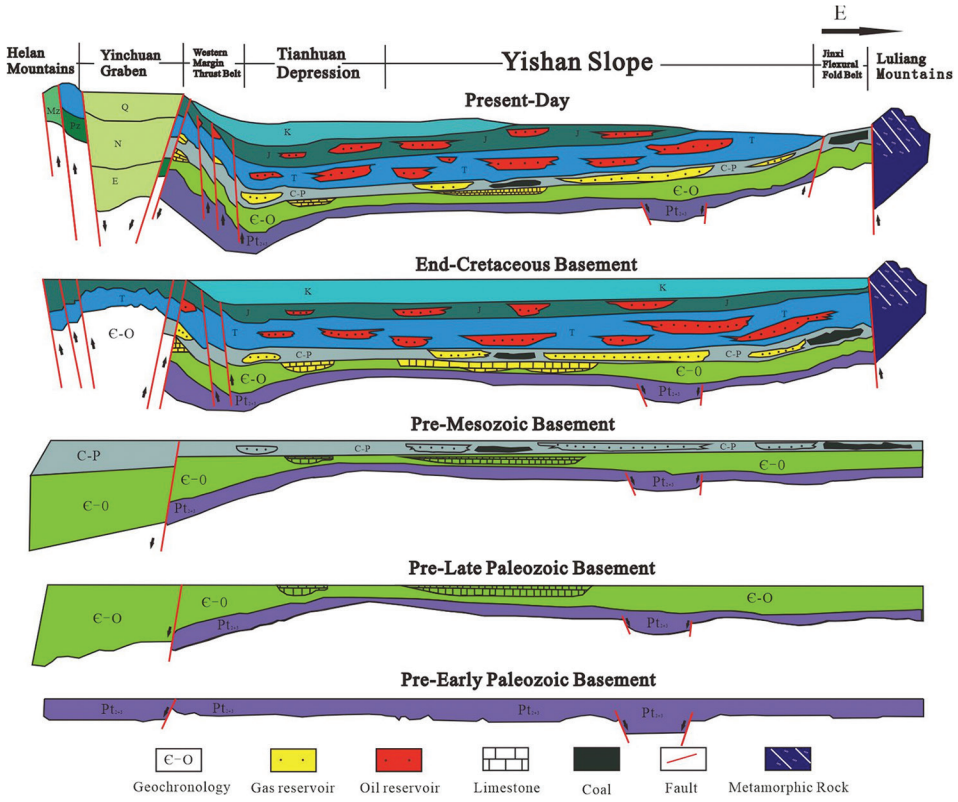



**Figure 11.** Structural sedimentary evolution profile of Ordos Basin. Adapted from Zhu *et al.*[27]

environments transitioned into delta-front terrigenous clastics, reflecting increasing continental influence. Toward the later stages, continued regression and shallowing resulted in the development of tidal-flat systems and widespread peat swamps. These conditions were highly favorable for the formation, preservation, and lateral continuity of thick and stable coalbeds (Figure 12). These long-lasting and relatively uniform depositional environments were essential for the extensive development of coal-bearing strata in the basin. The evolutionary sequence—from carbonate platform to clastic deltaic systems and finally to tidal-flat–peat complexes—provides key insights into the distribution, continuity, and internal heterogeneity of coalbeds, such as the No. 8 coalbed. Furthermore, the stability of the Yishan Slope ensured minimal tectonic disruption, allowing these coalbeds to maintain predictable structural patterns and making the region highly prospective for coal-related hydrocarbon resources.

The development of the No. 8 coal of the Benxi Formation in the YL area was controlled by the coupled effects of tectonics, climate, and sea-level change. During the Carboniferous period, the North China Plate was tectonically active, with peripheral orogeny inducing regional extension and facilitating epicontinental marine sedimentation. Under the background of Late Paleozoic glaciation, frequent glacial and interglacial cycles (1–2 Myr, with sea-level fluctuations of 50–100 m) periodically exposed coastal swamps, creating critical paleoclimatic conditions for extensive peat accumulation during the Benxi coal-forming period in North China.[29-31]

In the YL area, a key provenance reversal occurred between the Benxi 2 Formation and the Benxi 1 Formation. During the Benxi 2 Formation deposition, provenance was mainly derived from the northern Yinshan Uplift, resulting in quartzose sandstones associated with coalbeds. With the progressive uplift of the Qinling Orogenic Belt to the southeast, sediment supply during the Benxi 1 Formation shifted to the southeast, and depositional systems evolved into barrier coast and tidal-flat environments, favoring widespread development of organic-rich peat swamps. The prevailing warm and humid interglacial climate promoted lush vegetation growth and, coupled with slow subsidence, created favorable conditions for thick coalbed formation. High-frequency sea-level fluctuations induced dynamic hydrological changes, causing frequent vertical alternations of coal and mudstone, thereby generating coal gangue. Laterally, the No. 8 coal trends from southwest to northeast, thickening toward the northeast, with stable development of coalbeds up to 12 m thick in structural

lows[32-34] (Figure 13). Thus, this tectonic–climate–sea-level coupling mechanism establishes a crucial sedimentary facies model for seismic prediction of coal gangue distribution in the YL area of the Ordos Basin.

Variations in hydrodynamic energy directly control the probability and thickness of coal gangue development. In paleo-highland and paleo-slope areas, relatively strong hydrodynamic conditions favor the input of fine-grained clastic materials, resulting in frequent coal gangue occurrences within coal seams. In contrast, paleo-depressions are dominated by peat accumulation, where coal gangue development is relatively minimal. Sediment supply modes also exert a primary control on the spatial distribution of coal gangue. Line-source supply systems (e.g., deltaic or barrier island systems) tend to generate laterally continuous thin coal gangue layers, whereas point-source supply systems (e.g., localized channels) are more likely to produce discrete or clustered coal gangue distributions. High-frequency sea-level fluctuations, as discussed earlier, periodically modulate accommodation space and sediment supply, leading to repeated alternations between peat accumulation and clastic deposition, thereby significantly enhancing coal gangue development within specific paleogeomorphic units.

By projecting the seismic-predicted coal gangue probability volumes onto the reconstructed paleogeomorphological framework, the spatial probability distribution of coal gangue within the No. 8 coalbed was systematically mapped (Figure 14). The integrated results revealed a strong and consistent correlation between coal gangue development and paleogeomorphological units. Coal gangue-prone areas are predominantly concentrated in regions characterized by medium to high paleotopographic relief, where stronger hydrodynamic conditions and variable sediment supply favored the deposition of mixed lithofacies. In contrast, low-relief paleodepressions exhibit significantly lower coal gangue probabilities, reflecting a sedimentary environment dominated by relatively stable water conditions and continuous peat accumulation.

This geomorphologically constrained probability mapping not only visualizes 3D inter-well coal gangue distribution but also elucidates the relationship between coal gangue occurrence and paleogeomorphic evolution. These findings highlight the controlling role of paleotopography in governing sediment transport pathways, depositional environments, and coal gangue formation mechanisms. From an engineering perspective, the results provide direct and practical guidance for deep CBM development. During well pattern deployment, areas with high coal gangue probabilities should be avoided to maintain coalbed continuity and ensure sufficient reservoir thickness. For horizontal well trajectory design, integrating the predicted coal gangue distribution enables optimization of drilling paths, reducing the likelihood of intersecting coal gangue. This approach increases effective coal exposure, enhances horizontal section stability, and ultimately improves gas deliverability and production efficiency.

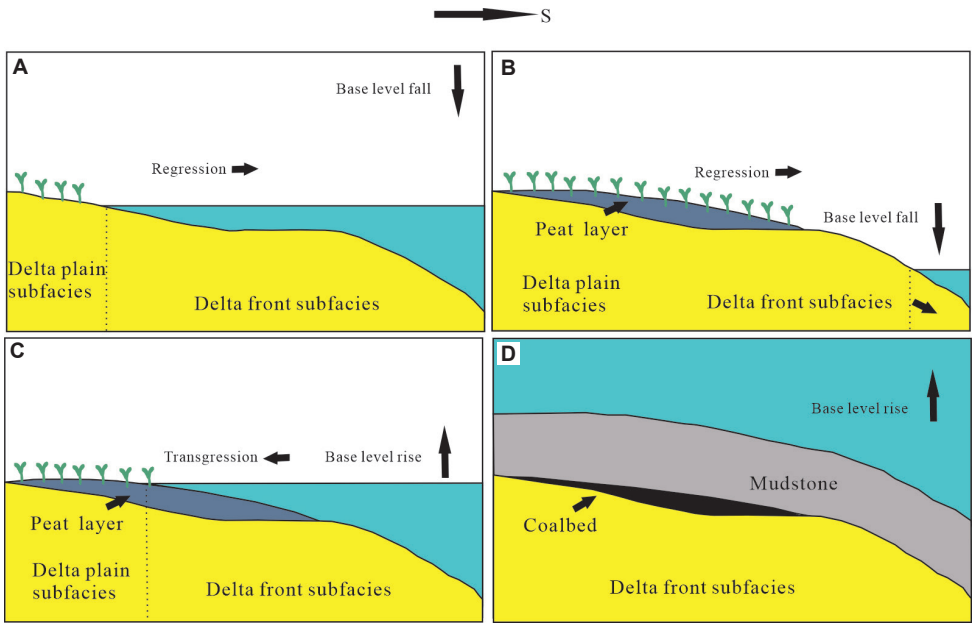



**Figure 12.** Coal accumulation model of the No. 8 coalbed in the Benxi Formation: (A) Early depositional stage of Mbr 1, Benxi FM (Late Carboniferous); (B) Late depositional stage of Mbr 1, Benxi FM (Late Carboniferous); (C) Late depositional stage of Mbr 1, Benxi FM (Late Carboniferous); (D) Depositional period of Taiyuan FM (Early Permian)

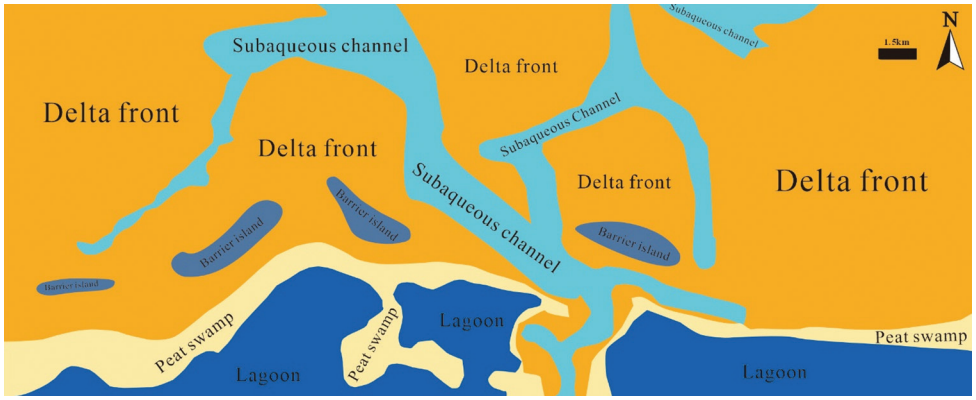



**Figure 13.** Sedimentary facies schematic diagram of the study area (before No. 8 coalbed deposition)

## 5. Conclusion

In predicting coal gangue, this study selected GR as the most sensitive log, and geostatistical stochastic optimization seismic inversion was performed to obtain the spatial probability volumes of coal gangue layers under paleogeomorphic zonation constraints. The predicted spatial distribution on maps, sections, and well-to-well profiles showed a high degree of consistency with log data, demonstrating strong geological reliability and engineering applicability. This confirms the robustness and practical value of the proposed method.

Based on geological, log, and geophysical data, the structural characteristics and sedimentary facies assemblages of the No. 8 coalbed in the Benxi Formation of the YL study area were clarified. The paleogeomorphology of the study area was identified as higher in the west and lower in the east, revealing a depositional evolution pattern in which the coalbed thickness gradually increases from west to east.

The geological features and spatial distribution of coal gangue in the YL area were clarified, and a strong coupling relationship between paleogeomorphology and coal gangue development was established, with an accuracy of 81.82%. Low-relief paleogeomorphological zones correspond to the lowest probability of coal gangue occurrence, making them favorable targets for coalbed development. Overall, by delineating the spatial distribution of coal gangue and interpreting its genetic and depositional controls, this study provides reliable theoretical support and technical guidance for the efficient exploration and development of deep CBM resources in the Ordos Basin.

## Acknowledgments

We thank PetroChina Changqing Oilfield Company for permitting us to use their log and seismic data in this study.

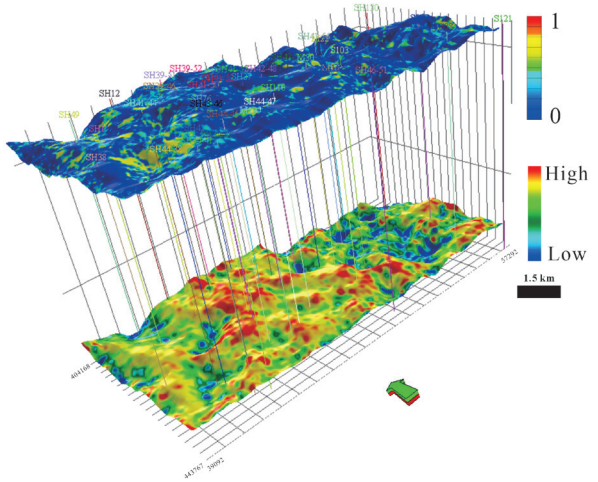



**Figure 14.** Three-dimensional display of coal gangue probability distribution projected onto paleogeomorphology

## Funding

This study was supported by the Sichuan Science and Technology Program (No. 2025ZNSFSC0004), the National Natural Science Foundation of China (No. 42241206), and the National Natural Science Foundation of China (No. U24B2022).

## Conflict of interest

The authors declare that they have no known financial or personal conflicts of interest that could have influenced the work reported in this paper.

## Author contributions

*Conceptualization:* Yucong Huang
*Data curation:* Yucong Huang
*Formal analysis:* Zelei Jiang
*Funding acquisition:* Xuri Huang
*Methodology:* Zelei Jiang
*Project administration:* Dong Zhang

*Resources:* Xuri Huang
*Supervision:* Xuri Huang, Dong Zhang, Yong Wu
*Validation:* Zelei Jiang
*Visualization:* Zelei Jiang
*Writing–original draft:* Zelei Jiang
*Writing–review & editing:* Zelei Jiang, Xuri Huang

## Availability of data

The data in this study are not openly available as they are confidential data provided by PetroChina Changqing Oilfield Company.

## References


1. Li G, Jia C, Zhao Q, Zhou T, Gao J. Coal-rock gas accumulation mechanism and the whole petroleum system of coal measures. *Pet Explor Dev*. 2025;52(1):33-49.

   doi: 10.1016/S1876-3804(25)60003-6

2. Li S, Wang C, Wang H, *et al*. Reservoir forming characteristics and favorable area evaluation of deep coalbed methane in Daning-Jixian Block. *Coal Geol Explor*. 2022;50(9):59-67.

   doi: 10.12363/issn.1001-1986.21.12.0842

3. Guo G, Liu Y, Wang C, Lv Y, Zhu X, Han G. Jiyu meiceng "san jiegou" de meiyan pinzhi zonghe pingjia [Comprehensive evaluation of coal rock quality based on "three structures" of coal seam]. *China Offshore Oil and Gas*. 2022;34(2):35-41.

   doi: 10.11935/j.issn.1673-1506.2022.02.004

4. Zhang L, Cai W. Mei cengqi chengxiang cejing ziliao zonghe pingjia fangfa [The evaluation method of imaging well logging data in coalbed methane reservoir]. *Well Log Technol.* 2009,33(6):571-574.

   doi: 10.3969/j.issn.1004-1338.2009.06.016

5. Sun B, Zeng F, Li M, Qi F. Xishan meitian malan kuangqu 8 hao mei jiqi jiagan de weiliang yu xitu yuansu diqiu huaxue tezhen [Geochemistry characteristics of trace elements and rare earth elements(REEs) of No. 8 coal and parting in Malan Coal Mine, Xishan Coalfield]. *J China Coal Soc*. 2010,35(01):110-116.

   doi: 10.13225/j.cnki.jccs.2010.01.002

6. Li Y. *Fuza Yingli Huanjing Han Jiagan Tehou Meiceng Dingmei Maoluo Guilv Ji Fangmei Gongyi Youhua [Caving Behavior of Top Coal and Optimization of Coal Drawing Technology in Extra-Thick Coal Seams with Gangue under Complex Stress Conditions*] China: China University of Mining and Technology; 2023.

7. Yan X, Xiong X, Xu F, *et al*. Shenbu meiyan chuceng dizhi lixue yingxiang jizhi ji kongzhi yinsu-yi eerduosi pendi daji qukuai weili [Deep coal geomechanical influence mechanism and its control factors of Daji Block in Ordos Basin]. *J China Coal Soc*. 2025;50(5):2550-2566.

   doi: 10.13225/j.cnki.jccs.2025.0037

8. Huang Q, Cheng J. Ruanying huceng yanti caichang diban de yingli fengbu ji pohuai tezheng yanjiu [Research on stress distribution and failure characteristics of coal mining floor in soft-hard alternant strata]. *Rock Soil Mech*. 2017,38(S1):36-42.

   doi: 10.16285/j.rsm.2017.s1.004

9. Gong F, Huang A, Kang W, *et al*. The influence of lamination and fracture on the velocity anisotropy of tectonic coals. *Geophysics*. 2024;89(6):MR355-MR365.

   doi: 10.1190/geo2024-0033.1

10. Gong F, Cheng J, Wang G, Peng S, Zhang Z. The effect of lamination on elastic anisotropy of primary coals under confining pressure: Experiment and theoretical modelling. *Geophys Prospect*. 2025;73(4):1228-1242.

    doi: 10.1111/1365-2478.70009

11. Wu K, Li S, Zhang H, *et al*. Initial characteristics of gas desorption in high-rank soft and hard coals and the causes of the differences. *ACS Omega*. 2025;10(30):33208-33219.

    doi: 10.1021/acsomega.5c03116

12. Wu K, Li S, Shi S, *et al*. A rapid method for the determination of coal seam gas pressure based on raw coal adsorption. *Energy Sci Eng*. 2025;13(10):4660-4673.

    doi: 10.1002/ese3.70258

13. Zhang L, Kan Z, Zhang C, Tian S, Zeng S, Hao D. Permeability and stress sensitivity of coals with different fracture directions under cyclic loading-unloading conditions: A case study of the Xutuan coal mine in Huaibei coalfield, China. *Rock Mech Rock Eng*. 2025;59(1):951-969.

    doi: 10.1007/s00603-025-04769-1

14. Huang H, Shi J, Chen G, *et al*. Eerduosi pendi daning-jixian qukuai shenceng bahaomei jiagan jiqi channeng guanxi fenxi [Analysis of the characteristics of interbedded gangue in the deep No. 8 coal seam of the Daning-Jixian block in the Ordos Basin and its relationship with production capacity]. *China Pet Explor*. 2025;30(4):108-119.

    doi: 10.3969/j.issn.1672-7703.2025.04.008

15. Yang Y, Zhang Y, Zeng Q. Research on coal gangue recognition based on multi-layer time domain feature processing and recognition features cross-optimal fusion. *Measurement*. 2022;204:112169.

    doi: 10.1016/j.measurement.2022.112169

16. Xu F, Nie Z, Sun W, *et al*. Eerduosi pendi dongyuan shenbu mei cengqi gaoxiao kaifa lilun jishu tixi [Theoretical and technological system for highly efficient development of deep coalbed methane in Daning-Jixian Block]. *J China Coal Soc*. 2024;49(1):528-544.

    doi: 10.13225/j.cnki.jccs.yh23.1290

17. Tang S, Tang D, Yang J, *et al*. Eerduosi pendi daning-jixian qukuai shenbu mei chuceng kongxi jiegou tezheng

ji chuceng qianli [Pore structure characteristics and gas storage potential of deep coal reservoirs in Daning-Jixian block of Ordos Basin]. *Acta Petrolei Sin*. 2023;44(11):1854-1866, 1902.

doi: 10.7623/syxb202311008

18. Xu F, Wang C, Xiong X, Xu B, Wang H, Zhao X. Eerduosi pendi dongyuan shenbu mei cengqi chengcang yanhua guilv yu kantan kaifa Shijian [Evolution law of deep coalbed methane reservoir formation and exploration and development practice in the eastern margin of Ordos Basin]. *Acta Petrolei Sin*. 2023;44(11):1764-1780.

doi: 10.7623/syxb202311002

19. Yan X, Xu F, Xiong X, *et al*. Shenbu mei cengqi kantan kaifa guanjian shiyan jishu ji fazhan fangxiang [Key experimental technologies and their development directions for the exploration and production of deep coalbed methane]. *Coal Geol Explor*. 2025;53(1):128-141.

doi: 10.12363/issn.1001-1986.25.01.0046

20. Zhang Y, Wu C, Li Y, *et al*. Coal-forming sedimentary model and its control on vertical reservoir heterogeneity of the upper carboniferous Benxi formation in Ordos basin, China. *Sci Rep*. 2025;15(1):28676.

doi: 10.1038/s41598-025-13928-x

21. Wang S, Duan L, Zhang M, Chen Z. Heterogeneity of coalbed methane pools and their exploration and development. *Pet Geol Exp*. 2000;22(4):368-370.

doi: 10.11781/sysydz200004368

22. Gong F, Zhang Z, Peng S, Guo Q, Wang G. Pre-stack seismic inversion of dual-porosity geometry in deep coalbed methane reservoirs based on decoupled equivalent medium theory. *Geophys Prospect*. 2026;74(1):e70131.

doi: 10.1111/1365-2478.70131

23. Wu Q, Xu H, Zou X. An effective method for 3D geological modeling with multi-source data integration. *Comput Geosci*. 2005;31(1):35-43.

doi: 10.1016/j.cageo.2004.09.005

24. Keith Turner A. Challenges and trends for geological modelling and visualisation. *Bull Eng Geol Environ*. 2006;65(2):109-127.

doi: 10.1007/s10064-005-0015-0

25. Yang Z, Lin C, Yin H, *et al*. Zhu chengfen fenxi zai Taizhong diqu aotaoxi yingshan zu tansuan yan yanxing shibie Zhong de yingyong [Application of principal component analysis in carbonate lithology identification of the Ordovician Yingshan Formation in Tazhong area]. *Nat Gas Geosci*. 2015;26(1):54-59.

doi: 10.11764/j.issn.1672-1926.2015.01.0054

26. Chong R, Yu X. Chuceng sanwei dizhi jianmo zhiliang kongzhi de guanjian dian [The keys to quality control of 3D reservoir modeling]. *Mar Geol Front*. 2011;27(7):64-69.

doi: 10.16028/j.1009-2722.2011.07.007

27. Zhu X, Deng X, Liu Z, Sun B, Liao J, Hui X. Daxing aoxian hupeng qianshui bianzhuang he sanjiaozhou chenji tezheng ji moshi: Yi eerduosi pendi longdong diqu yanchang zu weili [Sedimentary characteristics and model of shallow braided delta in large-scale lacustrine: An example from Triassic Yanchang Formation in Ordos Basin]. *Earth Sci Front*. 2013;20(2):19-28.

doi: cnki: sun:dxqy.0.2013-02-005

28. Wang J, Wang J. Low-amplitude structures and oil-gas enrichment on the Yishaan Slope, Ordos Basin. *Pet Explor Dev*. 2013;40(1):52-60.

doi: 10.1016/S1876-3804(13)60005-1

29. Zhao Z, Xu W, Zhao Z, *et al*. Geological characteristics and exploration breakthroughs of coal rock gas in carboniferous benxi formation, ordos basin, NW China. *Pet Explor Dev*. 2024;51(2):262-278.

doi: 10.1016/s1876-3804(24)60022-4

30. Cai Y, Gao G, Liu D, Qiu F. Eerduosi pendi dongyuan linxing zhongqu meixi qi fuji dizhi tiaojian ji chengzang moshi [Geological conditions for coal measure gas enrichment and accumulation models in Linxingzhong Block along the eastern margin of the Ordos Basin]. *Nat Gas Ind*. 2022;42(11):25-36.

doi: 10.3787/j.issn.1000-0976.2022.11.003

31. Jiang F, Jia C, Pang X, *et al*. Upper Paleozoic total petroleum system and geological model of natural gas enrichment in Ordos Basin, NW China. *Pet Explor Dev*. 2023;50(2):250-261.

doi: 10.1016/s1876-3804(23)60387-8

32. Li Y, Wu P, Gao J, *et al*. Mei cheng qi duoceng xi fuji jizhi yu quan hanqi xitong moshi-yi eerduosi pendi dongyuan linxing qukuai weili [Multilayer coal-derived gas enrichment mechanism and whole gas bearing system model: A case study on the Linxing Block along the eastern margin of the Ordos Basin]. *Nat Gas Ind*. 2022;42(6):52-64.

doi: 10.3787/j.issn.1000-0976.2022.06.005

33. Chen S, Tang D, Hou W, et al. Shenbu meiceng qi dizhi tiaojian teshu xing yu chuceng gongcheng xiangying [Geological particularity and reservoir engineering response of deep coalbed methane]. *Acta Petrolei Sin*. 2023;44(11):1993-2006.

doi: 10.7623/syxb202311018

34. Sun F, Zhou G, Tian W, Deng Z, Shen J. Meiceng qi xitong de dingyi\neihan\xingcheng ji yingyong-yi eerduosi pendi shitan xi-erdie xi meiceng weili [Definition, connotation, formation and application of coalbed methane system: A case study on the Carboniferous-Permian coal seams in the Ordos Basin]. *Nat Gas Ind*. 2024;44(7):42-53.

doi: 10.3787/j.issn.1000-0976.2024.07.004

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# TFDenoiser-Edge: A hybrid convolutional neural network–Transformer framework for real-time seismic denoising on edge devices in extreme environments

**Zesheng Yang[1,2], Qingfeng Xue[1,2]*, Xiaoning Wang[1,2], and Tao Wang[1,2]**

[1]Key Laboratory of Deep Petroleum Intelligent Exploration and Development, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China

[2]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China

## Abstract

Seismic monitoring in extreme environments, such as arid regions adjacent to the Alxa Desert, faces significant challenges due to complex noise interference from dust storms, high wind noise, and thermal variations. This paper presents TFDenoiser-Edge, a novel hybrid deep learning framework that combines convolutional neural networks (CNN) and Transformer architectures for real-time seismic signal denoising on resource-constrained edge devices. The proposed model employs a U-Net encoder–decoder structure with Transformer modules for global feature modeling in the time-frequency domain. To enable deployment on edge neural processing units (NPUs) with limited memory (≤512 MB), we introduced a mixed-precision quantization strategy that applies INT8 quantization to CNN layers while maintaining BF16 precision for Transformer layers, achieving 3.6× model compression with only 0.3 dB signal-to-noise ratio (SNR) loss. Additionally, a block-wise computation approach reduces peak memory consumption from 86 MB to 7.8 MB. Extensive experiments on Gansu seismic data demonstrated that TFDenoiser-Edge achieved an average SNR improvement of 8.5 dB, with P-wave and S-wave detection rates increasing from 65% to 91% and 52% to 85%, respectively. The model achieved real-time inference with 68 ms latency on edge NPUs while consuming less than 5 W of power, making it suitable for autonomous seismic monitoring in arid and desert regions. The proposed framework demonstrates potential generalizability to other extreme environments through transfer learning with minimal fine-tuning.

***Keywords:*** Seismic denoising; Edge computing; Transformer; Convolutional neural network; Mixed-precision quantization; Desert environment; Time-frequency analysis

***Corresponding author:**
Qingfeng Xue
(xueqingfeng@mail.iggcas.ac.cn)

**Citation:** Yang Z, Xue Q, Wang X, Wang T. TFDenoiser-Edge: A hybrid convolutional neural network–Transformer framework for real-time seismic denoising on edge devices in extreme environments. *J Seismic Explor.* 2026;35(2):025010138. doi: 10.36922/JSE025010138

**Received:** December 29, 2025

**Revised:** January 16, 2026

**Accepted:** January 26, 2026

**Published online:** March 12, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

In seismic observation and analysis, noise interference has long been a critical factor limiting data reliability and interpretation accuracy, particularly in extreme observation environments. Therefore, noise suppression is a core component of seismic data processing workflows. Over the past decades, researchers have proposed numerous

seismic data denoising methods, typically based on differences between the signal of interest and noise in the time or transform domains. These methods aim to improve data quality by enhancing the effective signal and suppressing noise. Typical time-domain denoising methods include polynomial fitting and median filtering.[1,2] However, these methods are sensitive to noise statistical characteristics, and their denoising performance is typically limited under low signal-to-noise ratio (SNR) conditions. In contrast, transform-domain denoising methods (such as those based on the Fourier, wavelet, S-, and curvelet transforms) utilize various mathematical transformations to convert seismic signals into the transform domain, separating effective signals from noise based on their differences in the transform domain. These methods have been widely applied in random noise attenuation and signal enhancement tasks.[3,4] However, these traditional methods typically rely on prior experience for parameter selection, and their performance is sensitive to noise type and data characteristics. In cases where the noise and signal spectra overlap substantially, or the SNR is low, it is often difficult to balance noise suppression and effective signal fidelity.[5]

In extreme observation environments, such as deserts and the Gobi regions, seismic data are often affected by strong winds, dust events, and diurnal temperature variations, resulting in environmental noise with distinct nonstationary and broadband characteristics. Related studies have shown that wind noise can introduce vibration interference covering a wide frequency band through surface coupling. Meanwhile, dust and temperature changes further exacerbate the overlap between noise and effective signals in spectral and spatiotemporal characteristics, making it difficult for traditional denoising methods based on fixed parameters or linear assumptions to effectively distinguish between signal and noise, leading to attenuation or distortion of seismic events.[6-8] Therefore, the applicability and robustness of traditional denoising methods still face significant challenges in extreme environments.

In recent years, deep learning methods have made significant progress in seismological applications,[9,10] often outperforming traditional signal processing methods. The combination of artificial intelligence (AI) and seismic monitoring systems has opened up new possibilities for geophysical forward modeling and inversion, automatic earthquake detection, phase picking, and signal enhancement.[11-16] Similarly, in seismic data denoising tasks, denoising models based on convolutional neural networks (CNNs), residual networks, and self-supervised learning frameworks can automatically learn the intrinsic feature representations of signals and noise through large-scale data training, thereby achieving effective noise suppression and signal enhancement without explicit prior parameter settings.[17-19] Although deep learning models have demonstrated excellent denoising performance in cloud environments, deploying complex models directly remains a significant challenge for real-time seismic monitoring in extreme environments, due to limitations in computational resources, power budgets, and the low-latency inference requirements of edge devices. Existing research has focused on optimizing deep learning models for edge inference scenarios using techniques such as quantization to achieve low-latency, energy-efficient inference under limited resources, which has become a core research topic.[20,21] This paper addresses these challenges by presenting TFDenoiser-Edge, a hybrid CNN–Transformer framework specifically designed for edge deployment in extreme environments. Our key contributions include: (i) A novel network architecture combining the local feature extraction capabilities of CNNs with the global modeling power of Transformers for effective time-frequency domain denoising; (ii) A mixed-precision quantization strategy that differentially handles CNN and Transformer components to balance compression ratio and accuracy; (iii) A block-wise computation approach that dramatically reduces memory footprint for edge deployment; and (iv) Comprehensive validation on Gansu seismic data demonstrating practical applicability in desert environments.

## 2. Related works

### 2.1. Deep learning for seismic signal processing

Deep learning has revolutionized seismic signal processing across multiple tasks. CNNs have been successfully applied to earthquake detection,[22,23] achieving high accuracy in distinguishing seismic events from noise. The introduction of attention mechanisms and Transformer architectures[24] has further enhanced model capabilities, particularly for capturing long-range dependencies in seismic waveforms. However, most existing models focus on accuracy optimization without considering deployment constraints on resource-limited devices.

### 2.2. Edge computing for seismic monitoring

Edge computing has emerged as a critical technology for real-time seismic monitoring systems. By processing data locally on edge devices, these systems can provide immediate responses without relying on network connectivity. However, the computational and memory constraints of edge devices[25,26]—typically featuring ARM processors with limited RAM and neural processing units (NPUs) with limited precision support—necessitate specialized model optimization techniques, including

quantization,[27] pruning, and efficient architecture design.[28]

## 3. Methodology

### 3.1. Network architecture overview

TFDenoiser-Edge adopts a U-Net-style[29] encoder–decoder architecture with Transformer modules integrated at the bottleneck layer for global modeling. The network processes three-component seismic data transformed into time-frequency representations via the short-time Fourier transform. The architecture consists of five key stages: input projection (Stage 1), hierarchical encoding (Stage 2), Transformer global modeling (Stage 3), hierarchical decoding (Stage 4), and mask generation (Stage 5). The overall network architecture is illustrated in Figure 1.

The input to the network is a time-frequency spectrogram with dimensions 256 × 256 × 6, where 256 × 256 represents the time-frequency resolution and 6 channels correspond to the real and imaginary parts of three-component seismic data. The output is a 256 × 256 × 1 mask that is element-wise multiplied with the noisy spectrogram to produce the denoised signal.

### 3.2. Input projection layer (Stage 1)

The input projection layer transforms the 6-channel input into a 32-channel feature representation through two consecutive convolutional layers, each followed by batch normalization and rectified linear unit (ReLU) activation. The convolution operation for the first layer is defined as:

$$Y(i,j,c) = \Sigma W(m,n,c',c) \cdot X(i+m, j+n, c') + b(c) \quad (1)$$

where $(i,j)$ denotes the spatial position, $c$ is the output channel, and $W$ and $b$ are learnable weights and biases, respectively. Batch normalization[30] stabilizes training by normalizing feature distributions within each mini-batch. The data flow through Stage 1 is shown in Figure 2.

### 3.3. Encoder (Stage 2)

The encoder employs a three-level downsampling structure, with each level containing two convolutional layers followed by max pooling. This design progressively expands the receptive field while extracting multi-scale features from fine to coarse granularity. The first encoder block processes the Stage 1 output and expands channels from 32 to 64:

$$E(1) = BN(Conv(Conv(Z(1), W_{conv1}), W_{conv2})) \quad (2)$$

Max pooling then reduces spatial dimensions by half. After three encoder blocks, the feature map dimensions are reduced from 256 × 256 × 32 to 64 × 64 × 128, achieving a 4× downsampling ratio while increasing channel depth from 32 to 128.

### 3.4. Transformer module (Stage 3)

The Transformer module addresses the limitation of CNNs' local receptive fields by modeling long-range dependencies across the entire time-frequency representation. This capability is crucial for identifying and suppressing persistent noise patterns such as sandstorm noise and prolonged wind noise specific to the Alxa Desert fringe.

The encoder output is first flattened into a sequence with length $L = 4096$ tokens. Learnable positional encodings are added to preserve spatial information. The multi-head self-attention mechanism[24] computes: $E(3) \in R^{64\times64\times128} X_{seq\downarrow}$

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

where $Q$, $K$, and $V$ are query, key, and value matrices obtained through linear projections. The model uses 4 attention heads with 16 dimensions per head. Four Transformer layers are stacked to progressively refine feature representations, balancing modeling capacity with computational efficiency for edge deployment. The

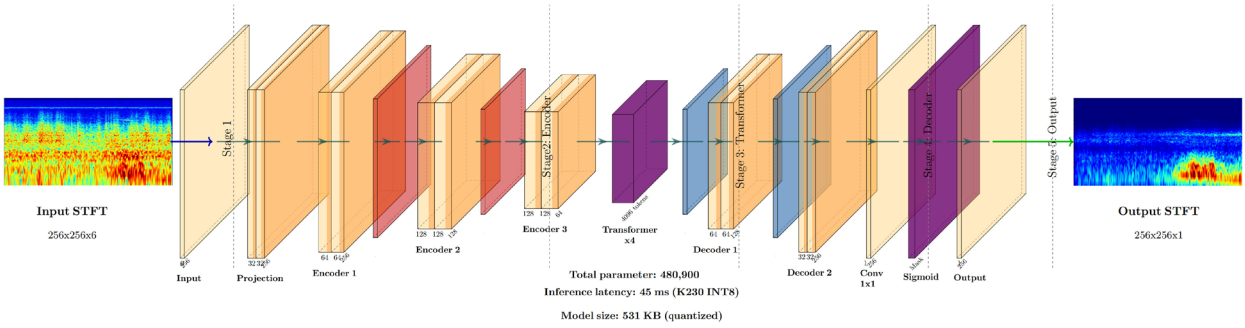



**Figure 1.** Network architecture. The entire network can be divided into five key stages: input projection (Stage 1), encoder (Stage 2), Transformer global modeling (Stage 3), decoder (Stage 4), and output generation (Stage 5).

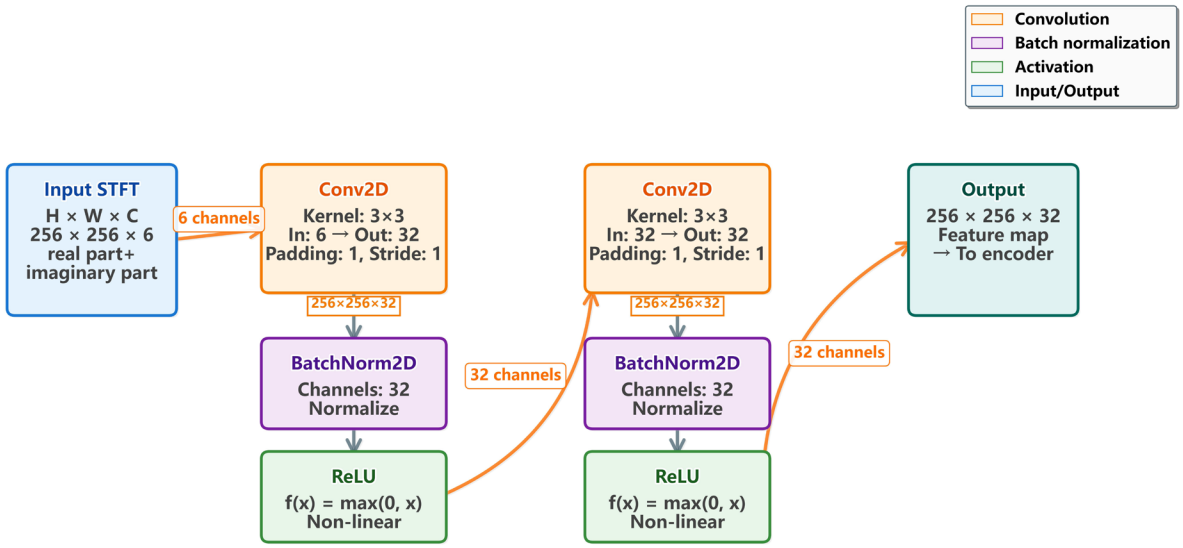



**Figure 2.** Stage 1 partial data flow. This layer consists of two consecutive convolutional layers, each followed by batch normalization and a rectified linear unit (ReLU) activation function. The second convolutional layer repeats the above process, mapping 32 channels to 32 channels, maintaining the feature dimensions.
Abbreviation: STFT: Short-time Fourier transform.

detailed structure of the Transformer module is depicted in Figure 3.

### 3.5. Decoder and mask generation (Stages 4-5)

The decoder progressively upsamples features back to the original 256 × 256 resolution through a symmetric three-level structure. Skip connections[31] from the encoder preserve fine-grained details by concatenating encoder features with upsampled decoder features. The final stage applies a 1 × 1 convolution followed by sigmoid activation to generate a mask $M \in [0,1]$:

$$M(i,j) = \sigma(Z_{\text{mask}}(i,j)) = \frac{1}{1+exp(-Z_{\text{mask}}(i,j))} \quad (4)$$

The denoised spectrogram is obtained by element-

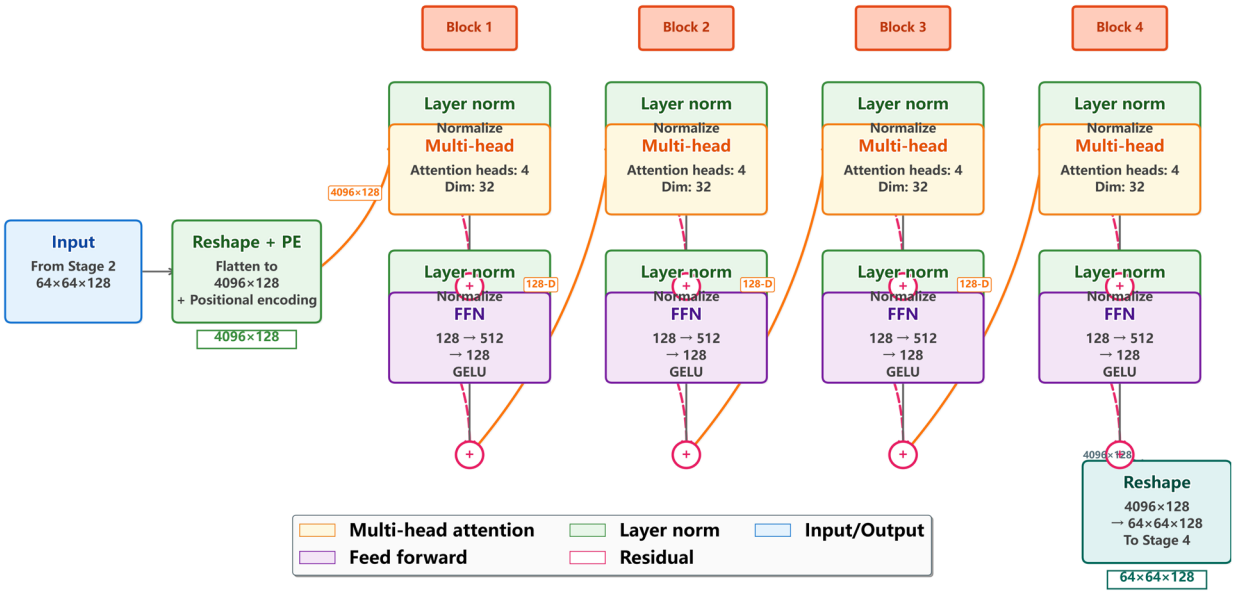



**Figure 3.** Detailed structural diagram of the Stage 3 Transformer module
Abbreviations: FFN: Feed forward network; GELU: Gaussian error linear unit; PE: Positional encoding.

wise multiplication: $\tilde{X}(i,j) = M(i,j) \odot X_{\text{noisy}}(i,j)$. This mask-based approach preserves phase information and generalizes well to unseen noise types.

### 3.6. Loss function

The model is trained with a multi-task loss combining time-domain, frequency-domain, and mask supervision:

$$L_{\text{total}} = \lambda_{\text{time}} L_{\text{time}} + \lambda_{\text{freq}} L_{\text{freq}} + \lambda_{\text{mask}} L_{\text{mask}} \quad (5)$$

where $L_{\text{time}}$ uses scale-invariant signal-to-distortion ratio (SI-SDR), $L_{\text{freq}}$ employs the magnitude spectrum mean squared error, and $L_{\text{mask}}$ applies binary cross-entropy. The weights are set as: $\lambda_{\text{time}} = 1.0$, $\lambda_{\text{freq}} = 0.5$, and $\lambda_{\text{mask}} = 0.3$.

For different datasets, the proposed framework primarily allows for parameter adjustments during data preprocessing and training. During the model adaptation phase, the model uses the AdamW (Adam with Weight Decay) optimizer,[32] which incorporates weight decay during parameter updates to improve the generalization.

## 4. Edge deployment optimization

### 4.1. Memory bottleneck analysis

Edge seismic monitoring devices typically feature ARM Cortex-A73 processors with 2.6 TOPS NPUs, 2 GB system memory, and strict power constraints (≤5 W). The available memory for inference is often limited to 512 MB or less, shared with the operating system and other applications. Deep neural networks consume memory primarily through: (i) model parameters, (ii) intermediate activations, and (iii) temporary buffers.

Analysis revealed that activation memory—not parameter storage—constitutes the primary bottleneck. For TFDenoiser-Edge at FP32 precision, the parameters require 124 MB, while the activations can exceed 391 MB during layer-by-layer computation. This far exceeds edge device capabilities.

### 4.2. Block-wise computation strategy

We proposed a block-wise computation approach that partitions the network into N blocks $\{B_1, B_2, \ldots, B_N\}$, processing one block at a time. This strategy exploits the insight that intermediate activations are only required for updating trainable parameters—frozen parameters do not require activation storage.

The approach implements: (i) Sequential block processing with storage of only boundary activations; (ii) Selective activation retention based on parameter update requirements; (iii) Memory recycling where block memory is immediately released after computation; and (iv) Dynamic memory allocation to prevent accumulation. This reduces peak memory from 86 MB to 7.8 MB—an 11× reduction. Figure 4 demonstrates how the proposed block-wise computation approach reduces the peak memory consumption.

### 4.3. Mixed-precision quantization strategy

Traditional uniform quantization to INT8[27] causes unacceptable accuracy degradation for Transformer layers. Our experiments revealed that CNN layers are robust to INT8 quantization (only 0.1 dB SNR loss), while Transformer layers suffer a 1.7 dB loss due to the sensitivity of the softmax exponential operations and residual connection error accumulation.

We proposed a mixed-precision strategy: CNN layers use INT8 symmetric quantization for maximum compression, while Transformer layers maintain BF16 precision to preserve accuracy. BF16 retains FP32's dynamic range (8-bit exponent) while reducing memory by 50%.[33] This achieved 3.6× overall compression with only 0.3 dB SNR loss, compared to 1.7 dB for full INT8 quantization. The mixed-precision quantization framework is shown in Figure 5, with performance comparison presented in Figure 6 and detailed sensitivity analysis in Table 1.

**Table 1. Quantization sensitivity analysis**

| Quantization scheme | SNR loss (dB) | Compression | Speedup |
|---|---|---|---|
| FP32 (baseline) | 0.0 | 1.0× | 1.0× |
| Full INT8 | 1.7 | 4.0× | 3.8× |
| Full BF16 | 0.1 | 2.0× | 1.2× |
| Mixed (proposed) | 0.3 | 3.6× | 2.8× |

Abbreviation: SNR: Signal-to-noise ratio.

## 5. Experiments

### 5.1. Dataset construction

The training dataset was constructed using three-component seismic data collected from Gansu monitoring stations in July 2024. Clean seismic events were synthetically combined with real background noise recordings at controlled SNR levels (0–20 dB) to create labeled training samples. The noise library encompassed dust-storm noise (40%), strong-wind noise (30%), thermal noise/instrumental noise (20%), and desert microseisms (10%), reflecting the actual noise distribution in arid environments adjacent to the Alxa Desert. Data augmentation techniques, including time shifting, amplitude scaling, and polarity reversal, were applied to enhance sample diversity. The final training set contained

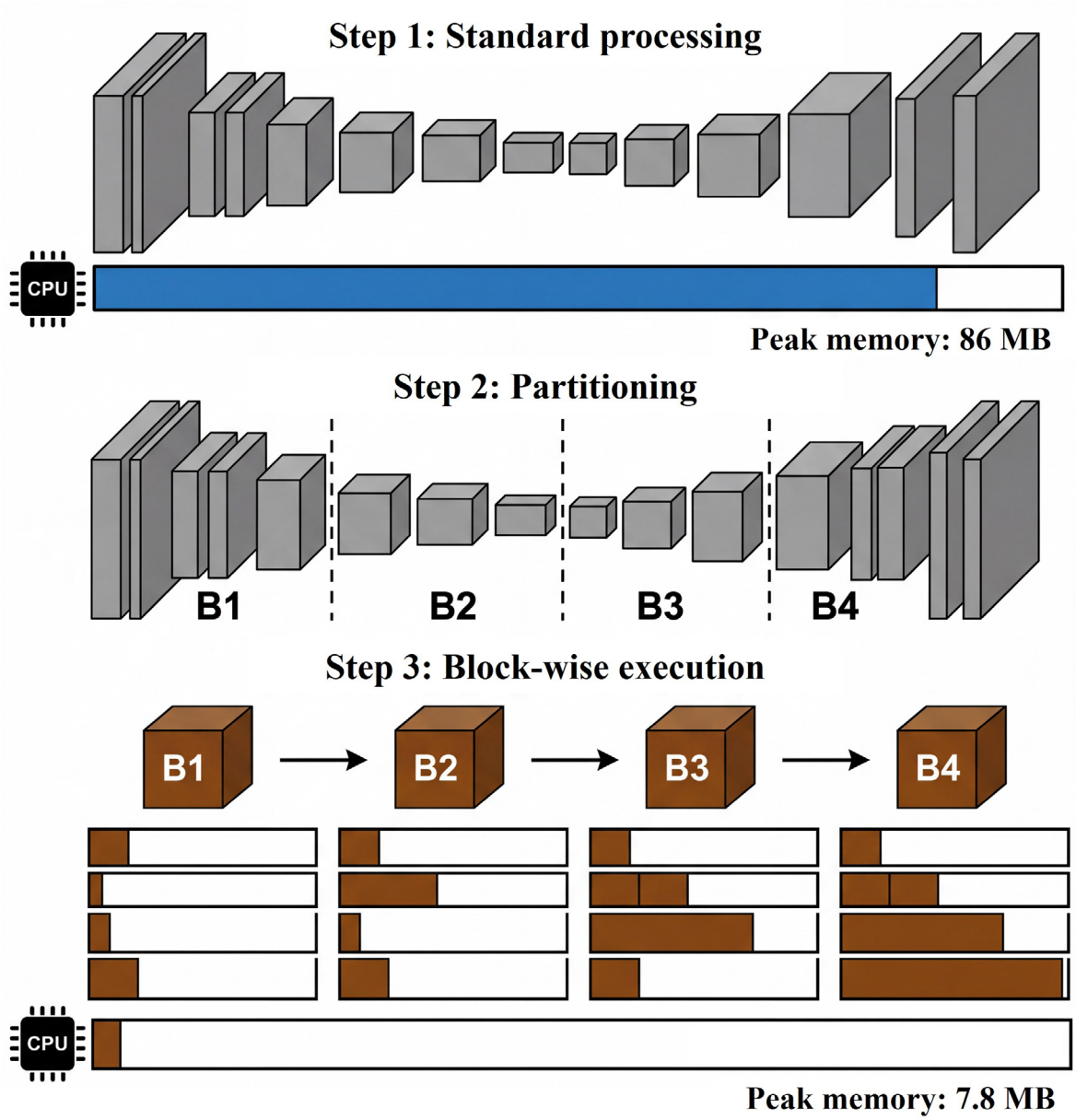



**Figure 4.** Comparison of peak memory consumption between standard full-model processing and the proposed block-wise partitioning approach

50,000 three-component samples. Representative training samples are visualized in Figure 7.

The validation set comprised 5,000 samples that were temporally and spatially independent from training data, including extremely low-SNR cases (<0 dB) and real recorded events with manual annotations. This setting ensures robust evaluation of generalization performance.

### 5.2. Evaluation metrics

Model performance was evaluated using: (i) SNR improvement—the increase in SNR after denoising; (ii) root mean square error (RMSE)—waveform reconstruction accuracy; (iii) structural similarity index (SSIM)—time-frequency structure preservation; (iv) peak SNR (PSNR)—a signal reconstruction quality metric; and (v) phase detection rate—accuracy of P-wave and S-wave arrival time picking using short-term average/long-term average (STA/LTA)[34] on denoised waveforms. All comparison methods used the same STA/LTA parameters. The equations for calculating these indicators are as follows:

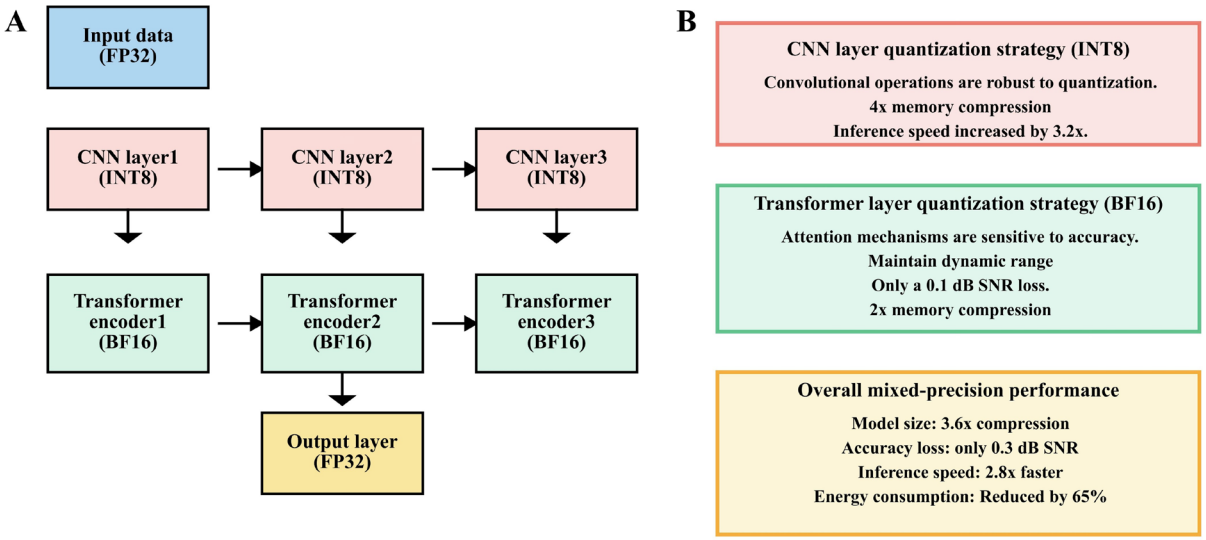



**Figure 5.** Framework diagram of mixed-precision quantization strategy. (A) The quantization process. (B) Detailed explanations of the technical characteristics and performance indicators of each quantization strategy.

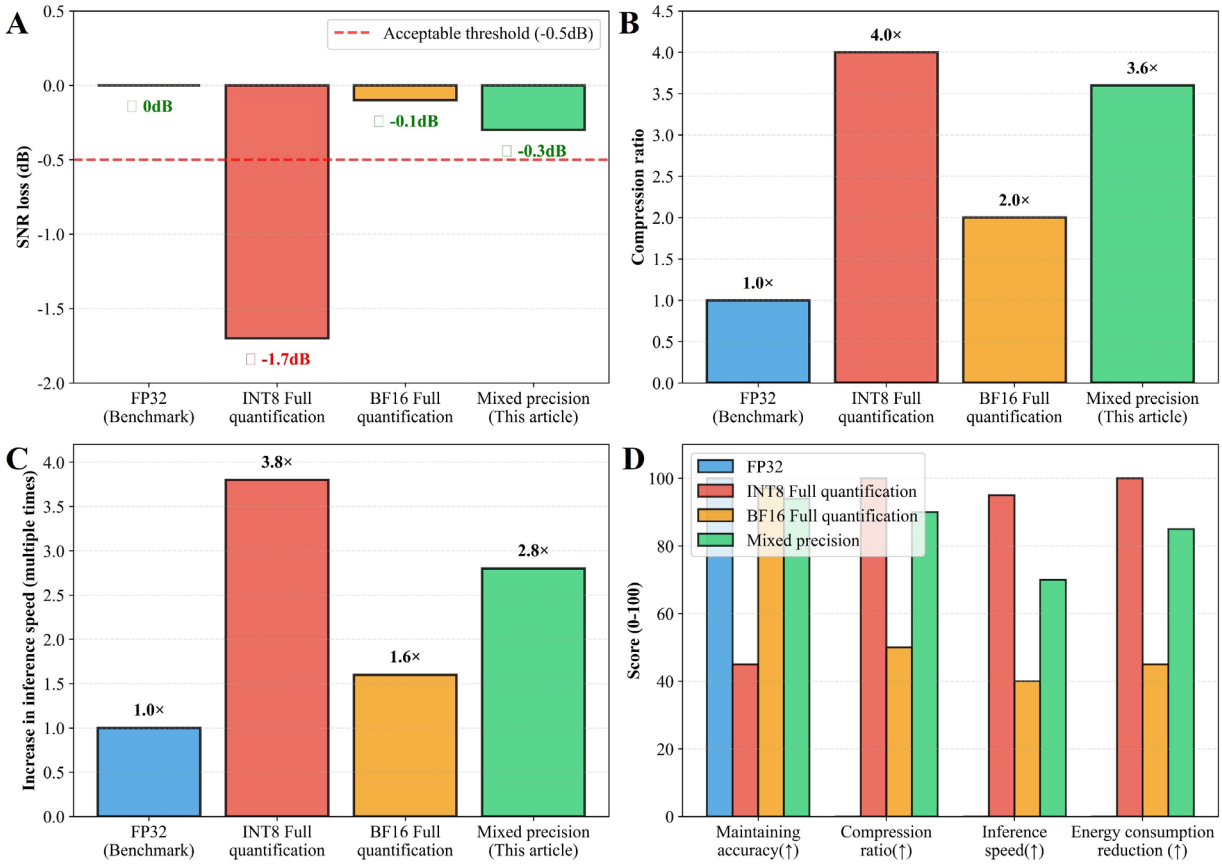



**Figure 6.** Performance comparison of different quantization schemes. (A) Signal-to-noise ratio (SNR) loss comparison. A green checkmark indicates acceptable performance, while a red cross indicates unacceptable loss. (B) Model compression ratio comparison. (C) Inference speed improvement comparison. (D) The comprehensive performance rating radar chart evaluates each solution across four dimensions: accuracy retention, compression ratio, inference speed, and power consumption reduction.

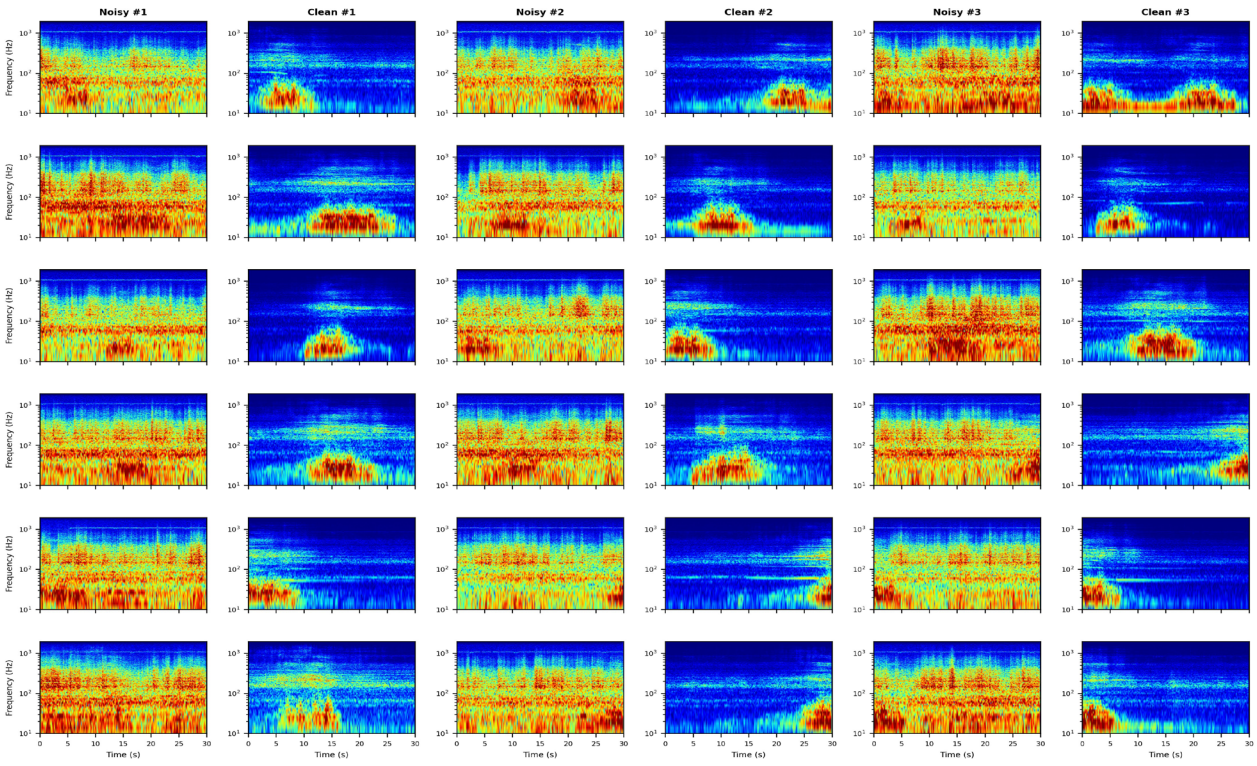



**Figure 7.** Visualization of training set data samples. Each row has a set of training samples, from left to right: Z-component noisy time-frequency map, Z-component clean event, N-component noisy time-frequency map, N-component clean event, E-component noisy time-frequency map, E-component clean event.

$$SNR_{improvement} = SNR_{out} - SNR_{in} \tag{6}$$

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\left(\hat{s}(n) - s(n)\right)^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\hat{s}(n) - s(n)\right)^2} \tag{8}$$

$$SSIM(X,Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \tag{9}$$

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

$$PDR = \frac{N_{correct}}{N_{total}} \times 100\% \tag{11}$$

where $SNR_{in}$ and $SNR_{out}$ represent the SNR before and after denoising, respectively; $\hat{s}(n)$ represents the reconstructed seismic waveform after denoising, $s(n)$ is the reference pure signal, and $N$ is the total number of sampling points; X and Y represent the time-frequency representations corresponding to the reference signal and the denoised signal, respectively, $\mu_X$ and $\mu_Y$ are their means, $\sigma_X$ and $\sigma_Y$ are their variances, $\sigma_{XY}$ is the covariance, and $C_1$ and $C_2$ are stable constants; $N_{total}$ is the total number of true P-wave or S-wave phases, and $N_{correct}$ is the number of phases correctly picked within a given tolerance $\Delta t$ ($|t_{pick} - t_{ref}| \leq \Delta t$ , where $t_{pick}$ is the picking time obtained based on the STA/LTA algorithm, and $t_{ref}$ is the true phase time in the manually labeled or synthetic data).

### 5.3. Results on synthetic data

Table 2 summarizes the quantitative results on the validation set. TFDenoiser-Edge achieved an average SNR improvement of 8.5 dB across all test conditions. The RMSE between denoised and clean waveforms was 0.15 (normalized), representing a 74% reduction from the noisy input RMSE of 0.58.

The SSIM of 0.87 indicates high structural similarity between predicted and target time-frequency masks. Component-wise analysis showed that the Z component achieved 0.89 SSIM, while both the N and E components achieved 0.86, demonstrating balanced performance across all channels. The PSNR increased from 12.3 dB to 28.5 dB, a 16.2 dB improvement.

Phase detection performance showed substantial improvements. P-wave detection rate increased from 65% to 91%, while S-wave detection improved from 52% to 85%—gains of 26 and 33 percentage points, respectively. The larger improvement for S-waves reflects their greater susceptibility to noise masking. Arrival time errors

decreased by 77% for both phases: P-wave error reduced from 0.35 s to 0.08 s, and S-wave error from 0.52 s to 0.12 s. Figure 8 visualizes the validation results.

**Table 2. Performance evaluation results**

| Metric | Before | After denoising |
|---|---|---|
| SNR improvement (dB) | - | +8.5 (average) |
| RMSE (normalized) | 0.58 | 0.15 |
| SSIM | - | 0.87 |
| PSNR (dB) | 12.3 | 28.5 |
| P-wave detection rate (%) | 65 | 91 |
| S-wave detection rate (%) | 52 | 85 |
| P-wave timing error (s) | 0.35 | 0.08 |
| S-wave timing error (s) | 0.52 | 0.12 |

Abbreviations: PSNR: Peak signal-to-noise ratio; RMSE: Root mean square error; SNR: Signal-to-noise ratio; SSIM: Structural similarity index.

### 5.4. Results on real Gansu data

The model was validated on actual seismic recordings from Gansu seismic stations during a continuous monitoring campaign (July 2024). Nineteen distributed intelligent seismic monitoring systems were deployed in the Gansu seismic network and surrounding desert areas. The edge AI module demonstrated robust performance under extreme conditions (high diurnal temperature swings and heavy dust), achieving an average inference latency of 68 ms and an SNR improvement of 8.2 dB.

Real data analysis confirms the model's ability to accurately identify seismic events in the time-frequency domain. The predicted event masks showed clear boundaries distinguishing signal from noise, with high-probability regions (mask values approaching 1) accurately localizing earthquake energy. The model effectively suppressed high-frequency dust-storm noise while preserving seismic signal content in the 2–3 Hz primary frequency band. Figures 9 and 10 present examples of real data processing results.

### 5.5. Edge deployment performance

Table 3 summarizes edge deployment characteristics. The optimized model achieved an average latency of 68 ms with P99 latency below 100 ms, enabling real-time processing

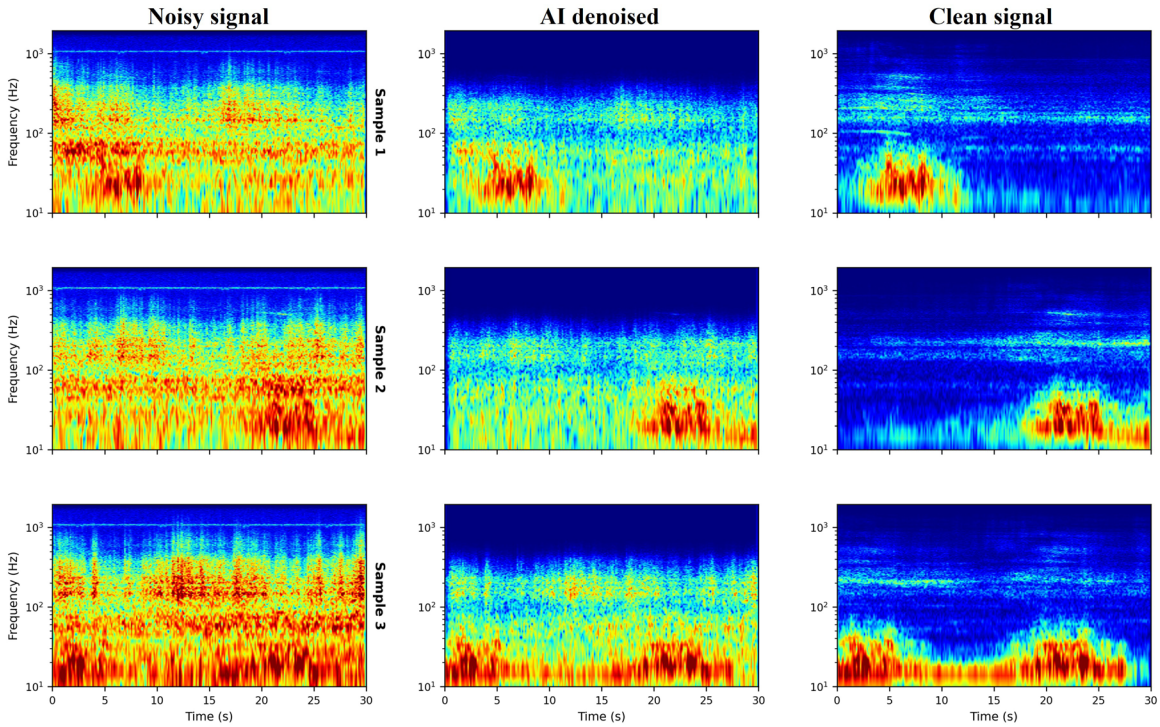



**Figure 8.** Visualization of validation set results. Each column corresponds to a component (Z, N, and E), from top to bottom: the input noisy time-frequency map, the model-predicted event mask, and the target event mask.
Abbreviation: AI: Artificial intelligence.

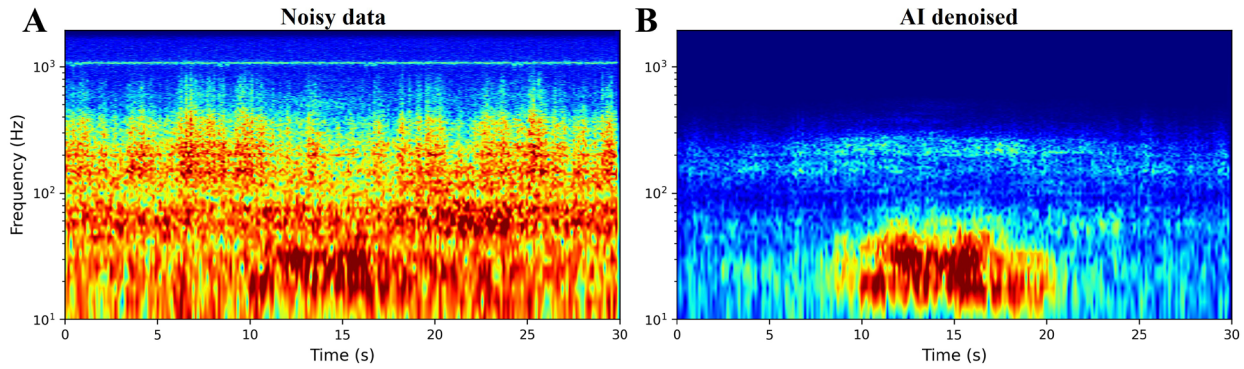



**Figure 9.** Processing results of the Z-component measured data (Example 1). (A) Input Z-component time-frequency plot (real part), with color coding representing normalized amplitude. (B) The clean event mask predicted by the model, with color-coded event probabilities (0–1).

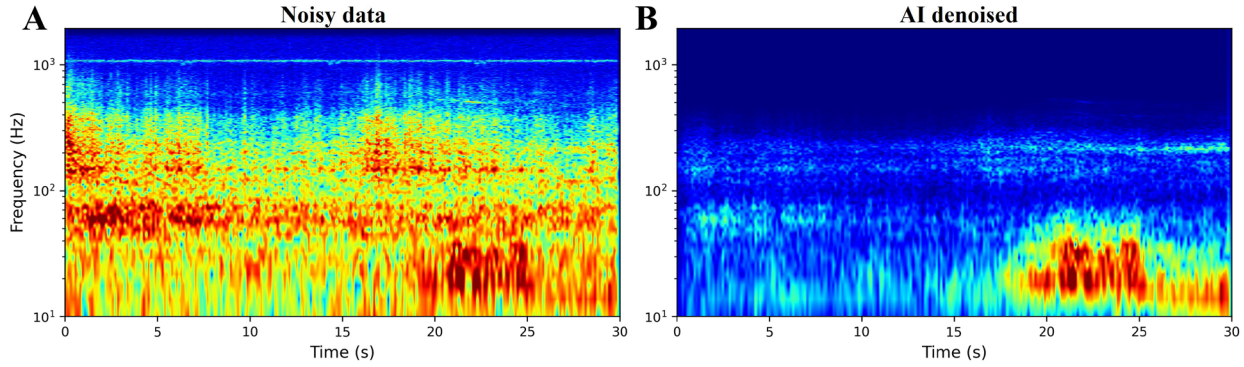



**Figure 10.** Processing results of the Z-component measured data (Example 2). (A) Input Z-component time-frequency plot (real part), with color coding representing normalized amplitude. (B) The clean event mask predicted by the model, with color-coded event probabilities (0–1).

of continuous seismic streams. Throughput reached approximately 15 samples/s. Power consumption remained below 5 W, meeting the requirements for battery-powered or solar-powered field deployment.

**Table 3. Edge deployment specifications**

| Specification | Value |
|---|---|
| Hardware platform | ARM Cortex-A73 + NPU (2.6 TOPS) |
| Model size (compressed) | 34.4 MB |
| Peak memory usage | 37 MB |
| Average latency | 68 ms |
| P99 latency | <100 ms |
| Throughput | ~15 samples/s |
| Power consumption | <5 W |
| Operating temperature | −40 °C to +85 °C |

### 5.6. Ablation study

To understand the contribution of each component, we conducted comprehensive ablation experiments on the validation set. The results are summarized in Table 4.

The Transformer module contributed most significantly to performance. Removing it reduced SNR improvement by 2.3 dB and SSIM by 0.09, confirming the importance of global modeling for distinguishing persistent noise patterns from transient seismic signals. The multi-head attention mechanism enables the model to capture long-range dependencies that CNNs alone cannot efficiently model.

Skip connections preserve fine-grained temporal details essential for accurate phase timing. Without them, the model showed 1.4 dB degradation in SNR improvement and a notable decrease in SSIM, indicating loss of structural information during the encoding-decoding process.

The multi-task loss function provided complementary supervision signals. Removing frequency-domain and

mask losses reduced performance by 0.7 dB, suggesting that joint optimization in multiple domains yields more robust feature learning.

Regarding Transformer depth, 2 layers showed suboptimal performance while 8 layers provided marginal improvement over 4 layers at a significantly higher computational cost. The four-layer configuration strikes an optimal balance between modeling capacity and edge deployment efficiency.

**Table 4. Ablation study results**

| Configuration | SNR improvement (dB) | SSIM |
|---|---|---|
| Full model (proposed) | 8.5 | 0.87 |
| w/o Transformer (CNN only) | 6.2 | 0.78 |
| w/o Skip connections | 7.1 | 0.82 |
| w/o Multi-task loss | 7.8 | 0.84 |
| 2 Transformer layers | 7.9 | 0.85 |
| 8 Transformer layers | 8.6 | 0.87 |

Abbreviations: CNN: Convolutional neural network; SNR: Signal-to-noise ratio; SSIM: Structural similarity index.

### 5.7. Comparison with baseline methods

We compared TFDenoiser-Edge with several baseline methods on the Gansu validation set, with results presented in Table 5: (i) Traditional bandpass filtering (1–10 Hz); (ii) Wiener filtering; (iii) Wavelet denoising with soft thresholding; (iv) DeepDenoiser[35,36]—a CNN-based denoising model; and (v) WaveDecompNet[37]—a U-Net-style network for seismic separation.

**Table 5. Comparison with baseline methods**

| Method | SNR (dB) | P-Det (%) | S-Det (%) |
|---|---|---|---|
| Bandpass filter | 2.1 | 68 | 55 |
| Wiener filter | 3.4 | 72 | 61 |
| Wavelet denoising | 4.2 | 75 | 64 |
| DeepDenoiser | 6.8 | 84 | 76 |
| WaveDecompNet | 7.3 | 86 | 79 |
| TFDenoiser-Edge | 8.5 | 91 | 85 |

Abbreviation: Det: Detection rate; SNR: Signal-to-noise ratio.

Traditional signal processing methods (Bandpass filter, Wiener filter, and Wavelet denoising) showed limited effectiveness against the complex noise patterns in arid desert environments. These methods assume specific noise characteristics that do not hold for the mixture of dust storm, wind, and instrumental noise present in desert recordings.

Deep learning methods (DeepDenoiser and WaveDecompNet)[35-37] outperformed traditional approaches by learning noise patterns directly from data. However, TFDenoiser-Edge achieved the best results due to its hybrid architecture that combines local feature extraction with global context modeling. The improvement was particularly notable for S-wave detection (6–9% higher than other deep learning methods), where distinguishing overlapping P-coda from S-wave onset requires global temporal context.

### 5.8. Field deployment experience

From July 18–23, 2024, we deployed three-component seismic nodes in the desert region of Gansu Province for a microseismic monitoring experiment. The complete system underwent rigorous field testing during the Gansu monitoring campaign, operating continuously under desert conditions.

Key deployment specifications included: (i) Actual operating temperature range: −10 °C to +40°C (observed field conditions during deployment); (ii) Power source: solar panels with battery backup; (iii) Communication: 4G cellular and Wi-Fi links when available; (iv) Storage: 128 GB local storage with automatic rotation; and (v) Maintenance interval: zero maintenance during the expedition period.

The edge AI module processed over 50,000 channel-hours of continuous three-component recordings. System logs confirmed a consistent average inference latency of 68 ms throughout the deployment period. No system failures or performance degradation were observed under harsh environmental conditions, validating the robustness of both the hardware design and the software optimization.

Real-time denoising enabled detection of 127 confirmed seismic events during the campaign, including 89 tectonic earthquakes, 31 regional events, and 7 events of uncertain origin. Without edge denoising, preliminary analysis suggests only 78 of these events would have been reliably detected using traditional methods—a 63% increase in detection capability attributed to the TFDenoiser-Edge system.

## 6. Discussion

The success of TFDenoiser-Edge stems from several key design decisions. The hybrid CNN–Transformer architecture[24,29,31] leverages complementary strengths: CNNs efficiently extract local time-frequency patterns while Transformers capture global dependencies necessary for distinguishing persistent noise from transient seismic

signals. The U-Net structure with skip connections preserves fine-grained details essential for accurate phase timing.

The mixed-precision quantization strategy addresses a fundamental challenge in deploying Transformer models on edge devices. The sensitivity analysis revealed that Transformer components are particularly vulnerable to aggressive quantization due to the exponential operations in softmax attention and error accumulation through residual connections. By selectively applying appropriate precision levels, we achieved near-lossless compression while maintaining practical deployment requirements.

The block-wise computation approach represents a novel solution to the activation memory bottleneck. Unlike gradient checkpointing, which trades memory for computation by recomputing activations during backpropagation, our method is optimized for inference scenarios where selective activation management based on frozen-parameter identification dramatically reduces peak memory usage without computational overhead.

Field deployment in a desert environment validates the practical viability of edge AI for extreme-environment monitoring. The system operated reliably under severe wind and temperature swing conditions, demonstrating the robustness of both the hardware design and the software optimization. The significant improvements in phase detection rates directly translate into enhanced earthquake location accuracy[38]—a 0.2 s reduction in timing error can improve location precision by 2–3 km.

Limitations and future work include: (i) The current model is trained on Gansu noise characteristics and may require transfer learning for other environments; (ii) Multi-station array processing could further improve detection capabilities; and (iii) Integration with real-time earthquake early warning systems remains to be developed.

### 6.1. Theoretical analysis of quantization sensitivity

The differential quantization sensitivity between CNN and Transformer layers can be explained through mathematical analysis. For the softmax attention mechanism,[12] small perturbations in the input are exponentially amplified:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{12}$$

where $\text{softmax}(x_i) = \frac{exp(x_i)}{\sum_j exp}$. The exponential function *exp(x)* amplifies quantization errors non-linearly. For example, when $x$ changes from 10.0 to 10.1 (1% change), $exp(x)$ changes from 22,026 to 24,343 (10.5% change). This sensitivity propagates through the attention computation, causing significant accuracy degradation under aggressive quantization.

Furthermore, Transformer's residual connections accumulate quantization errors across layers. For an L-layer Transformer with per-layer quantization error $\varepsilon_l$, the total accumulated error is approximately $\approx \sum_{l=1}^{L} \varepsilon_l$. With four layers and approximately 0.4 dB error per layer under INT8, the total approaches 1.6 dB—consistent with our experimental observation of 1.7 dB loss.

In contrast, CNN operations involve primarily linear transformations (convolutions) and piecewise-linear activations (ReLU). These operations are inherently more robust to quantization since small input perturbations produce proportionally small output changes without exponential amplification.

### 6.2. Practical considerations for deployment

Several practical factors influenced our deployment strategy. First, the choice of BF16 for Transformer layers rather than FP16 was motivated by dynamic range considerations. FP16's limited 5-bit exponent restricts dynamic range to $\pm 6.5 \times 10^4$, which can cause overflow in attention score computation. BF16's 8-bit exponent matches FP32's range[33] ($\pm 3.4 \times 10^{38}$), preventing numerical issues while halving memory requirements.

Second, the block-wise computation strategy required careful implementation to avoid performance degradation. Memory allocation and deallocation must be synchronized with computation to prevent fragmentation. We implemented a memory pool that pre-allocates buffers for all blocks and manages them through a lightweight scheduler, achieving consistent latency across inference iterations.

Third, protection against extreme heat and dust ingress proved critical in desert deployment. The edge computing module generates heat during intensive NPU operations while facing high ambient temperatures. Our hardware design includes enhanced heat dissipation optimized for high-temperature environments and appropriate sealing/filtration against fine sand, maintaining stable operating temperatures between −20 °C and +85 °C at the die level.

### 6.3. Comparison with cloud-based processing

While cloud-based processing can leverage more powerful hardware, edge deployment offers critical advantages for desert monitoring: (i) Latency—edge processing achieves 68 ms vs. seconds for round-trip cloud communication; (ii) Connectivity—desert stations often have intermittent or low-bandwidth communication links; (iii) Power—transmitting raw data consumes significantly more power than local processing; and (iv) Autonomy—edge systems continue operating during communication outages.

Our hybrid approach combines edge preprocessing with selective cloud upload. High-confidence detections (threshold >0.95) use edge results directly, while uncertain cases (0.5–0.95) are queued for cloud verification when connectivity permits. This strategy reduces data transmission by approximately 85% while maintaining detection reliability.

### 6.4. Generalization to other extreme environments

Although trained on Gansu desert data, the TFDenoiser-Edge framework is designed for generalization to other extreme environments. Time-frequency domain processing captures noise characteristics independent of specific geographic contexts. Preliminary tests on seismic data from polar environments and volcanic regions (characterized by tremor and explosion signals) showed promising results with minimal fine-tuning.[39,40]

For deployment in new environments, we recommend a transfer learning approach: (i) collect 2–4 weeks of continuous noise recordings from the target site; (ii) fine-tune only the decoder layers while freezing encoder and Transformer weights; and (iii) validate on held-out local data before deployment. This approach reduces adaptation time from weeks of full training to approximately 2–3 hours on the edge device itself.

## 7. Conclusion

This paper presents TFDenoiser-Edge, a hybrid CNN–Transformer framework[24,29] for real-time seismic denoising on edge devices deployed in extreme environments. The key contributions include a novel network architecture combining local and global feature modeling in the time-frequency domain, a mixed-precision quantization strategy differentially optimizing CNN and Transformer components, and a block-wise computation approach for memory-efficient edge deployment.

Extensive experiments demonstrated that TFDenoiser-Edge achieved 8.5 dB average SNR improvement while increasing P-wave and S-wave detection rates by 26 and 33 percentage points, respectively. The model runs in real time (68 ms latency) on edge NPUs, consuming less than 5 W of power. Successful deployment during the Gansu monitoring campaign validates the system's practical applicability for autonomous seismic monitoring in arid and desert regions.

This work represents the first successful deployment of Transformer-based deep learning models on seismic monitoring edge NPUs, providing both theoretical insights and practical solutions for intelligent geophysical sensing in resource-constrained environments. The proposed techniques are generalizable to other edge AI applications requiring efficient deployment of hybrid architectures.

## Acknowledgments

None.

## Funding

This work was supported by the National Key R&D Program of China (grant no. 2021YFA0716800), and the CAS Project for Young Scientists in Basic Research (grant no. YSBR-020).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Qingfeng Xue
*Formal analysis:* Zesheng Yang, Qingfeng Xue, Tao Wang, Xiaoning Wang
*Funding acquisition:* Qingfeng Xue
*Investigation:* Zesheng Yang, Qingfeng Xue, Tao Wang
*Methodology:* Zesheng Yang, Qingfeng Xue
*Validation:* Qingfeng Xue
*Visualization:* Zesheng Yang, Qingfeng Xue, Xiaoning Wang
*Writing–original draft:* Zesheng Yang
*Writing–review & editing:* Qingfeng Xue

## Availability of data

Data collected through the research presented in the paper are available upon request to the corresponding author.

## References

1. Liu G, Chen X, Li J, Du J, Song J. Seismic noise attenuation using nonstationary polynomial fitting. *Appl Geophys*. 2011;8(1):18-26.

   doi: 10.1007/s11770-010-0244-2

2. Wu S, Wang Y, Di Z, Chang X. Random noise attenuation by 3D Multi-directional vector median filter. *J Appl Geophys*. 2018;159:277-284.

   doi: 10.1016/j.jappgeo.2018.09.021

3. Mousavi SM, Langston CA. Hybrid Seismic Denoising Using Higher-Order Statistics and Improved Wavelet Block Thresholding. *B Seismol Soc Am*. 2016;106(4):1380-1393.

   doi: 10.1785/0120150345

4. Yu Z, Abma R, Etgen J, Sullivan C. Attenuation of noise and simultaneous source interference using wavelet denoising. *Geophysics*. 2017;82(3):V179-V190.

   doi: 10.1190/geo2016-0240.1

5. Shao J, Wang Y, Liang X, Xue Qi, Liang E, Shi S. *Ji yu luan*

*sheng wang luo de ren gong zhen yuan fen bu shi guang xian chuan gan shu ju zao sheng ya zhi* [Siamese network based noise elimination of artificial seismic data recorded by distributed fiber-optic acoustic sensing]. *Chin J Geophys*. 2022;65(9):3599-3609. [In Chinese].

doi: 10.6038/cjg2022P0919

6. Li Y, Yu W, Zhang C, Yang B. Low-frequency noise suppression for desert seismic data based on a wide inference network. *J Geophys Eng*. 2019;16(4):801-810.

doi: 10.1093/jge/gxz051

7. Zhang S, Li Y. Seismic exploration desert noise suppression based on complete ensemble empirical mode decomposition with adaptive noise. *J Appl Geophys*. 2020;180:104055.

doi: 10.1016/j.jappgeo.2020.104055

8. Zhong T, Ye Y. MFIEN: Multi-scale feature interactive enhancement network for seismic data denoising in desert areas. *Sci Rep*. 2025;15:3979.

doi: 10.1038/s41598-025-87481-y

9. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.

doi: 10.1038/nature14539

10. Mousavi SM, Beroza GC. Deep-learning seismology. *Science*. 2022;377(6607):eabm4470.

doi: 10.1126/science.abm4470

11. Zhu W, Beroza GC. PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method. *Geophys J Int*. 2019;216(1):261-273.

doi: 10.1093/gji/ggy423

12. Mousavi SM, Ellsworth WL, Zhu W, Chuang LY, Beroza GC. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat Commun*. 2020;11:3952.

doi: 10.1038/s41467-020-17591-w

13. Zheng Y, Wang Y, Liang X, *et al*. A deep learning approach for signal identification in the fluid injection process during hydraulic fracturing using distributed acoustic sensing data. *Front Earth Sci*. 2022;10:999530.

doi: 10.3389/feart.2022.999530

14. Zheng Y, Wang Y. Ground-penetrating radar wavefield simulation via physics-informed neural network solver. *Geophysics*. 2023;88(2):KS47-KS57.

doi: 10.1190/geo2022-0293.1

15. Wu S, Wang Y, Liang X. Joint denoising and classification network: Application to microseismic event detection in hydraulic fracturing distributed acoustic sensing monitoring. *Geophysics*. 2023;88(4):L53-L63.

doi: 10.1190/geo2022-0296.1

16. Li S, Yang X, Cao A, *et al*. SeisT: A foundational deep learning model for earthquake monitoring tasks. *IEEE Trans Geosci Remote Sens*. 2024;62:1-15.

doi: 10.1109/TGRS.2024.3371503

17. Shao J, Wang Y, Yao Y, Wu S, Xue Q, Chang X. Simultaneous denoising of multicomponent microseismic data by joint sparse representation with dictionary learning. *Geophysics*. 2019;84(5):KS155-KS172.

doi: 10.1190/geo2018-0512.1

18. Zhu W, Mousavi SM, Beroza GC. Seismic Signal Denoising and Decomposition Using Deep Neural Networks. *IEEE Trans Geosci Remote Sens*. 2019;57(11):9476-9488.

doi: 10.1109/TGRS.2019.2926772

19. Quinones L, Tibi R. Denoising Seismic Waveforms Using a Wavelet-Transform-Based Machine-Learning Method. *B Seismol Soc Am*. 2024;114(4):1777-1788.

doi: 10.1785/0120230304

20. Dantas PV, Sabino Da Silva W, Cordeiro LC, Carvalho CB. A comprehensive review of model compression techniques in machine learning. *Appl Intell*. 2024;54(22):11804-11844.

doi: 10.1007/s10489-024-05747-w

21. Ngo D, Park HC, Kang B. Edge Intelligence: A Review of Deep Neural Network Inference in Resource-Limited Environments. *Electronics*. 2025;14(12):2495.

doi: 10.3390/electronics14122495

22. Ross ZE, Meier MA, Hauksson E, Heaton TH. Generalized Seismic Phase Detection with Deep Learning. *B Seismol Soc Am*. 2018;108(5A):2894-2901.

doi: 10.1785/0120180080

23. Perol T, Gharbi M, Denolle M. Convolutional neural network for earthquake detection and location. *Sci Adv*. 2018;4(2):e1700578.

doi: 10.1126/sciadv.1700578

24. Vaswani A, Shazeer N, Parmar N, *et al*. Attention Is All You Need. *arXiv*. Preprint posted online 2017.

doi: 10.48550/arXiv.1706.03762

25. Seemakhupt K, Liu S, Khan S. EdgeRAG: Online-Indexed RAG for Edge Devices. *arXiv*. Preprint posted online 2024.

doi: 10.48550/arXiv.2412.21023

26. Zhang F, Zhang C, Guan J, *et al*. Breaking the Edge: Enabling Efficient Neural Network Inference on Integrated Edge Devices. *IEEE Trans Cloud Comput*. 2025;13(2):694-710.

doi: 10.1109/TCC.2025.3559346

27. Jacob B, Kligys S, Chen B, *et al*. Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference. In: *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition;* June 18-23

2018; Salt Lake City, UT, USA; 2018:2704-2713.

doi: 10.1109/cvpr.2018.00286

28. Howard AG, Zhu M, *et al*. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv*. Preprint posted online 2017.

doi: 10.48550/arXiv.1704.04861

29. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Proceedings of the Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Vol 9351*. Lecture Notes in Computer Science. MICCAI 2015; October 5-9 2015; Munich, Germany: Springer International Publishing; 2015:234-241.

doi: 10.1007/978-3-319-24574-4_28

30. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In: Bach F, Blei D, eds. *Proceedings of the 32nd International Conference on Machine Learning*. Volume 37: International Conference on Machine Learning; July 7-9 2015; Lille, France; 2015:448-456. https://proceedings.mlr.press/v37/ioffe15.html

31. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30 2016; Las Vegas, NV, USA; 2016:770-778.

doi: 10.1109/cvpr.2016.90

32. Zhou P, Xie X, Lin Z, Yan S. Towards Understanding Convergence and Generalization of AdamW. *IEEE Trans Pattern Anal Mach Intell*. 2024;46(9):6486-6493.

doi: 10.1109/TPAMI.2024.3382294

33. Micikevicius P, Narang S, Alben J, *et al*. Mixed Precision Training. In: *Proceedings of the 2018 International Conference on Learning Representations*. 6th International Conference on Learning Representations; April 30-May 3 2018; Vancouver, Canada; 2018. https://openreview.net/forum?id=r1gs9JgRZ

34. Allen RV. Automatic earthquake recognition and timing from single traces. *B Seismol Soc Am*. 1978;68(5):1521-1532.

doi: 10.1785/BSSA0680051521

35. Saad OM, Chen Y. Deep denoising autoencoder for seismic random noise attenuation. *Geophysics*. 2020;85(4):V367-V376.

doi: 10.1190/geo2019-0468.1

36. Yang L, Liu X, Zhu W, Zhao L, Beroza GC. Toward improved urban earthquake monitoring through deep-learning-based noise suppression. *Sci Adv*. 2022;8(15):eabl3564.

doi: 10.1126/sciadv.abl3564

37. Yin J, Denolle MA, He B. A multitask encoder–decoder to separate earthquake and ambient noise signal in seismograms. *Geophys J Int*. 2022;231(3):1806-1822.

doi: 10.1093/gji/ggac290

38. Lomax A, Michelini A, Curtis A. Earthquake Location, Direct, Global-Search Methods. In: Meyers, R, ed. *Encyclopedia of Complexity and Systems Science*. New York, NY, USA: Springer New York; 2009:2449-2473.

doi: 10.1007/978-0-387-30440-3_150

39. Lapins S, Butcher A, Kendall JM, *et al*. DAS-N2N: machine learning distributed acoustic sensing (DAS) signal denoising without clean data. *Geophys J Int*. 2023;236(2):1026–1041.

doi: 10.1093/gji/ggad460

40. Fernández-Carabantes J, Titos M, D'Auria L, García J, García L, Benítez C. RNN-DAS: A New Deep Learning Approach for Detection and Real-Time Monitoring of Volcano-Tectonic Events Using Distributed Acoustic Sensing. *JGR Solid Earth*. 2025;130(9):e2025JB031756.

doi: 10.1029/2025JB031756

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Critical metal enrichment in coal-bearing strata of the eastern Ordos Basin: A geophysical and geochemical study

**Yi Yang[1], Jingtao Zhao[2]*, Suzhen Shi[1], and Wanli Gao[1]**

[1]College of Geoscience and Survey Engineering, China University of Mining and Technology (Beijing), Beijing, China

[2]College of Water Sciences, Beijing Normal University, Beijing, China

## Abstract

Coal-bearing strata hold significant potential for critical metal resources. This study investigates their enrichment mechanisms in the eastern Ordos Basin using an integrated geophysical-geochemical approach. We combine seismic impedance inversion, tectonic evolution, and log facies interpretation with geochemical data (total organic carbon, sulfur, vitrinite reflectance, trace/rare earth elements) to reconstruct the sedimentary-tectonic-hydrogeological evolution. Results show notable enrichment of zirconium, chromium (Cr), nickel, and hafnium. Seismic facies indicate that deltaic distributary channels and tidal sandbars were primary conduits for metal-bearing fluids. The Taiyuan Formation was deposited in a warm-humid, freshwater-influenced tidal delta-barrier island system under reducing conditions, enabling initial metal sequestration via sulfidation and organic complexation. Reconstruction of tectonic evolution confirms that Yanshanian tectono-thermal events associated with the Zijinshan pluton generated fault networks that channeled hydrothermal fluids, driving secondary metal enrichment. Groundwater circulation, controlled by sedimentary facies and fractures, further regulates metal remobilization. A focused study on Cr delineates a three-stage enrichment model: source weathering, sedimentary reduction, and tectonic-hydrothermal activation. This work establishes a structure–sedimentation–groundwater coupled joint ore-controlling model, underscoring the value of integrated geophysical and geochemical methods for exploring critical metals in coal-bearing strata.

***Keywords:*** Coal measures; Critical metals; Seismic impedance inversion; Enrichment mechanisms; Sedimentary facies; Tectonic-thermal evolution; Groundwater circulation

***Corresponding author:**
Jingtao Zhao
(diffzjt@bnu.edu.cn)

**Citation:** Yang Y, Zhao J, Shi S, Gao W. Critical metal enrichment in coal-bearing strata of the eastern Ordos Basin: A geophysical and geochemical study. *J Seismic Explor*. 2026;35(2):025520130. doi: 10.36922/JSE025520130

**Received:** December 24, 2025

**Revised:** February 10, 2026

**Accepted:** February 13, 2026

**Published online:** March 27, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Coal-bearing strata are important reservoirs for strategic metals (e.g., gallium [Ga], germanium [Ge], lithium [Li]), offering significant economic potential beyond their conventional use as energy resources.[1,2] In the eastern margin of the Ordos Basin, previous research has primarily focused on coal accumulation and coalbed methane, leaving the enrichment mechanisms of critical metals within the coal measures poorly understood.

Geological structures, such as faults and folds, provide essential pathways and space

for metal-bearing hydrothermal fluids, controlling their migration, focusing, and eventual deposition.[3,4] These tectonic features, including faults, folds, and magmatic bodies, fundamentally influence ore-forming space, permeability, and fluid pathways, thereby governing metal transport and precipitation.[5–7] Multi-source hydrothermal fluids (magmatic, alkaline, volcanic) can mobilize and precipitate metals such as gold (Au) and Ge through interaction with sulfides and carbonates. Volcanic fluids, characterized by low pH, high volatile content, and significant ore-forming elements, act as efficient agents for leaching metals from wall rocks and precipitating them as sulfides or oxides upon cooling or mixing with groundwater.[8] Tectono-magmatic activities regulate both the source and intensity of mineralization[9], with magmatic intrusions enhancing basement leaching[10] and introducing exogenous metals via volcanic ash.[11] Fracture networks serve as primary conduits, controlling both primary enrichment and subsequent reactivation of metals.[12] Regional tectonic evolution exerts long-term control on metallogenesis. For instance, the Caledonian uplift provided structural conditions for Ga enrichment, while Yanshanian tectono-thermal events drove secondary metal enrichment through thermal alteration.[13] Techniques such as layer flattening, balanced sections, and thickness recovery have demonstrated how paleo-tectonic patterns regulate sediment supply and diagenetic fluid flow, ultimately shaping the three-dimensional (3D) distribution of metals in coal measures.[14–16]

Sedimentary environments also play a critical role, with redox conditions, provenance, and lithology collectively governing initial metal concentration and subsequent transformation.[17] In the northern Ordos Basin, the Carboniferous–Permian marine–terrestrial transitional setting established strongly reducing conditions that promoted organic matter adsorption and volcanic ash incorporation, leading to rare earth element (REE) enrichment—evident from characteristic negative europium (Eu) and cerium (Ce) anomalies in coals.[18] Conversely, Jurassic terrestrial oxidation inhibited element preservation. Sedimentary environments further influence syndepositional precipitation and lithofacies control on epigenetic uranium (U) mineralization.[19,20] In the Taiyuan Formation, alternating warm-humid and dry-hot phases favored freshwater anoxic conditions conducive to Ga enrichment. Advances in seismic stratigraphy and stratigraphic slicing now enable quantitative reconstruction of deltaic systems, clarifying how sandbody architecture and microfacies control metal migration.[21–24]

In contrast, the role of groundwater in trace metal enrichment remains understudied.[25–27] Metals such as Ga, selenium, strontium (Sr), zirconium (Zr), REEs, mercury, lead (Pb), and thorium (Th) are often mobilized and enriched through dissolution-reprecipitation during diagenesis and redox interface fluctuations.[28] For example, groundwater leaching has been proposed as a key factor in REE enrichment in Late Permian coal seams in the Sichuan Basin.[29]

In the study area, thick, cyclic coal measures with diverse source-reservoir-cap assemblages present a complex setting.[30–34] While seismic inversion and lithofacies modeling have begun to clarify tectono-sedimentary controls on 3D metal zoning, a comprehensive understanding of the synergistic effects of tectonic, sedimentary, and fluid systems—particularly in the eastern Ordos Basin—is still lacking. Moreover, the contribution of groundwater circulation to metal transport has not been adequately addressed. This study introduces an integrated framework that synergistically combines high-resolution seismic interpretation, geochemical provenance analysis, and paleogroundwater modeling to elucidate the multi-stage enrichment mechanisms of critical metals. Our approach advances the field by: (i) Integrating seismic facies analysis with trace element geochemistry to constrain the sedimentary environment's control on initial metal enrichment; (ii) Reconstructing the Late Paleozoic–Mesozoic tectonic uplift history to clarify the influence of tectono-magmatic activity on metal migration and re-concentration; and (iii) Establishing a structure–sedimentation–groundwater joint control model to provide a predictive framework for metal enrichment in analogous basin settings.

## 2. Geological setting

The study area is located within the Jinxi flexural fold belt on the eastern margin of the Ordos Basin (Figure 1). The structural pattern is characterized by a north–south orientation as the dominant feature, with faults present in the Paleozoic, yet their scale is limited.[35] In the southern section, there is an evident tectonic and magmatic superposition, and the faults display a semi-circular radial pattern associated with the Zijinshan pluton. Conversely, the northern region remains a slightly deformed area, unaffected by magmatism. The stratigraphic sequence comprises the Carboniferous Benxi Formation, the Taiyuan Formation, the Permian Shanxi Formation, the Lower Shihezi Formation, the Upper Shihezi Formation, and the Shiqianfeng Formation. The primary coal-bearing strata are the Taiyuan Formation and the Shanxi Formation.

## 3. Materials and methods

The LX-5 well, which is strategically situated in the

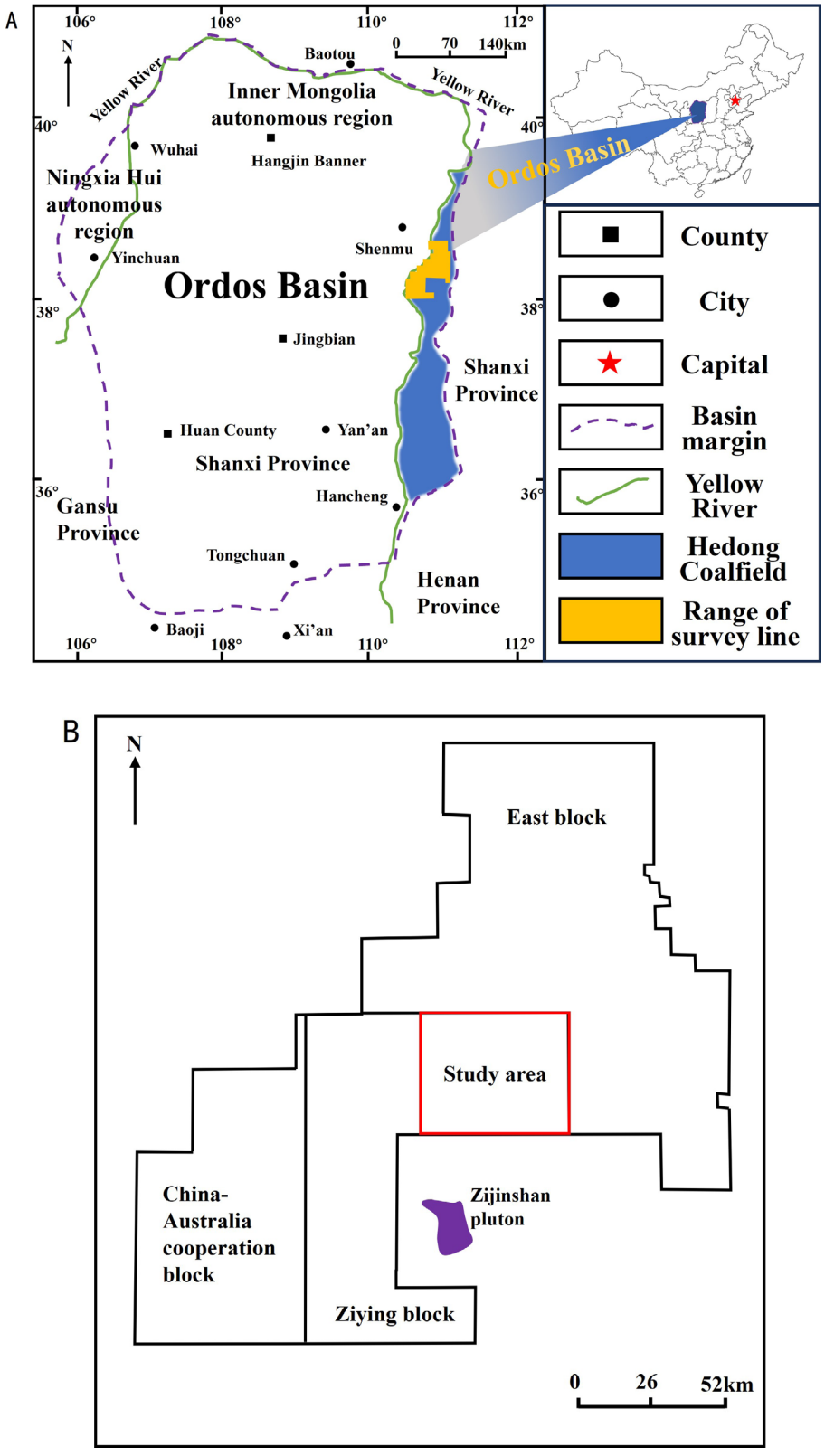



**Figure 1.** Location of the study area. (A) Geographical location of the Ordos Basin. (B) Distribution of the study area.

southwestern region of the study area (Figure 1B), penetrated the complete and continuous Carboniferous–Permian coal-bearing stratigraphic sequence. To ensure data representativeness and capture the geological variations related to the major depositional and tectonic stages, a systematic, high-density sampling methodology was employed along the entire cored section. Fifty-five core samples were collected at precise depths, encompassing all the major formations (Benxi, Taiyuan, Shanxi, Lower Shihezi, Upper Shihezi, and Shiqianfeng) and their predominant lithologies (coal, mudstone, and sandstone) (Table 1). Subsequently, a comprehensive series of analyses was conducted on this representative sample collection, including total organic carbon (TOC), sulfur (S) content, vitrinite reflectance, major and trace elements, and clay minerals. This stratified and systematic approach ensured that the resulting high-resolution vertical dataset effectively captures the key temporal sequences of processes necessary to investigate the mechanisms governing critical metal enrichment.

All experiments were conducted in the Laboratory of Mineralization and Dynamics at Chang'an University. The analysis and determination of TOC and S content were conducted using the carbon-S analyzer (CS-230, LECO Corporation, USA). Vitrinite reflectance was measured using a microphotometer (MPV-SP, Leica, Germany). Trace elements and REE were analyzed using an inductively coupled plasma mass spectrometer (X-series, Thermo Fisher, USA). Quantitative analysis of the mineral composition was conducted using an X-ray diffractometer (D/max-2200 2035C 4, Rigaku Corporation, Japan).

## 4. Results and discussion

### 4.1. Geochemical experiment

#### *4.1.1. Total organic carbon, sulfur content, and vitrinite reflectance*

To assess organic matter richness and paleoenvironmental conditions, we measured TOC and S content. Elevated S concentrations often signify anoxic depositional settings, facilitating organic preservation and sulfide formation. Analyses revealed considerable heterogeneity in organic abundance across the sequence. For instance, the Benxi Formation yielded high TOC values (avg. 27.72%), whereas the Upper Shihezi Formation was markedly lean (avg. 0.09%) (Table 1).

#### *4.1.2. Trace, rare earth elements, and mineral composition analysis*

The concentrations of trace elements and REE serve as the foundation for a comprehensive exploration of sediment provenance, sedimentation, and diagenesis, as well as for assessing the sedimentary environment. Figure 2 demonstrates that the contents of elements such as Sr, barium (Ba), chromium (Cr), and Pb exhibit significant fluctuations across different formations. The elements with the highest average content are Zr (216.46 μg/g), Ba (164.14 μg/g), and Sr (127.38 μg/g). In the study area, we analyzed the distribution of element content in Carboniferous–Permian strata. Figure 2A illustrates the content distribution of trace elements, where the contents of Ba and Sr are relatively high in the Carboniferous strata, which might be associated with the redox conditions of the sedimentary environment. Figure 2B depicts the REE distribution, showing distinct enrichment characteristics that may reflect changes in sediment provenance and the sedimentary environment.

To gain a comprehensive understanding of the mineral composition of the Upper Shihezi Formation and Benxi Formation, a mineralogical analysis was conducted. Figure 3A depicts the overall mineral composition of the sample, where quartz and clay minerals constitute the primary mineral components. Figure 3B illustrates the relative distribution of clay minerals. Illite and kaolinite are the predominant clay minerals, which implies a sedimentary environment ranging from semi-arid to humid.

### 4.2. Depositional environment

#### *4.2.1. Sedimentary facies control*

The sedimentary facies structure of coal-bearing strata was systematically studied by combining single well facies, cross-section sedimentary facies and seismic constrained plane facies. By establishing a quantitative relationship between well-derived microfacies and optimized seismic attributes (especially RMS amplitude, sweetness, and instantaneous phase), a comprehensive sedimentary model was constructed to reveal a unique delta and barrier island sedimentary system with logging characteristics, sand body geometry, and spatial phase distribution.

(i) Single well facies analysis: The integration of well-log motifs and core lithology from LX-5 facilitated the identification of a deltaic depositional system within the Shanxi Formation. Characteristic features included bell-shaped gamma-ray responses in the mudstones and coals of interdistributary bays, as well as box-shaped profiles associated with distributary channel sandstones. Sedimentary microfacies displayed distinctive log responses. Bell-shaped curves were manifested in the sequences of dark mudstone and coal within interdistributary bays. Deposits of distributary channels, which consisted of fine sandstone, medium sandstone, and pebbly sandstone, generate box-shaped and bell-shaped

**Table 1. Geochemical experimental analysis of the Shihezi Formation and Benxi Formation in LX-5 well**

| Samples | Depth (m) | Formation | Lithology | TOC (weight %) | S (weight %) | Vitrinite reflectance (R0, %) |
|---|---|---|---|---|---|---|
| L5-1-1 | 1,285.58 | Upper Shihezi | Mudstone | 0.09 | 0.01 | 1.09 |
| L5-1-3 | 1,619.57 | Lower Shihezi | Mudstone | 0.65 | 0.01 | 0.96 |
| L5-1-7 | 1,735.41 | Taiyuan | Mudstone | 7.69 | 0.90 | 0.92 |
| L5-1-8 | 1,736.58 | Taiyuan | Carbonaceous mudstone | 7.29 | 0.31 | 0.94 |
| L5-1-12 | 1,784.74 | Benxi | Mudstone | 1.13 | 0.56 | 1.65 |
| L5-1-11 | 1,784.74 | Benxi | Coal | 54.32 | 0.43 | 0.98 |

Abbreviations: S: Sulfur; TOC: Total organic carbon.

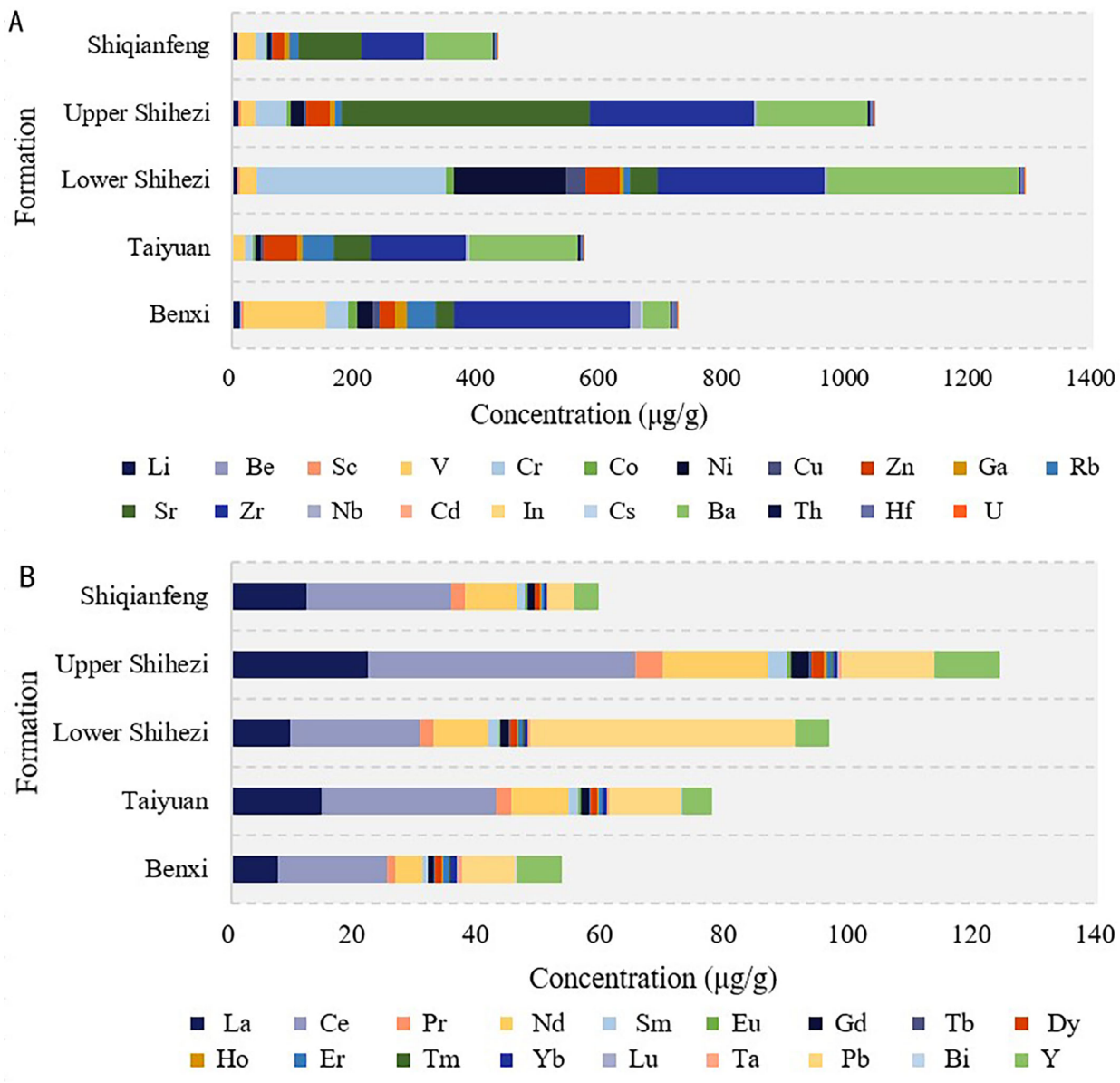



**Figure 2.** Element content distribution of the Shiqianfeng Formation–Benxi Formation. (A) Trace elements. (B) Rare earth elements.

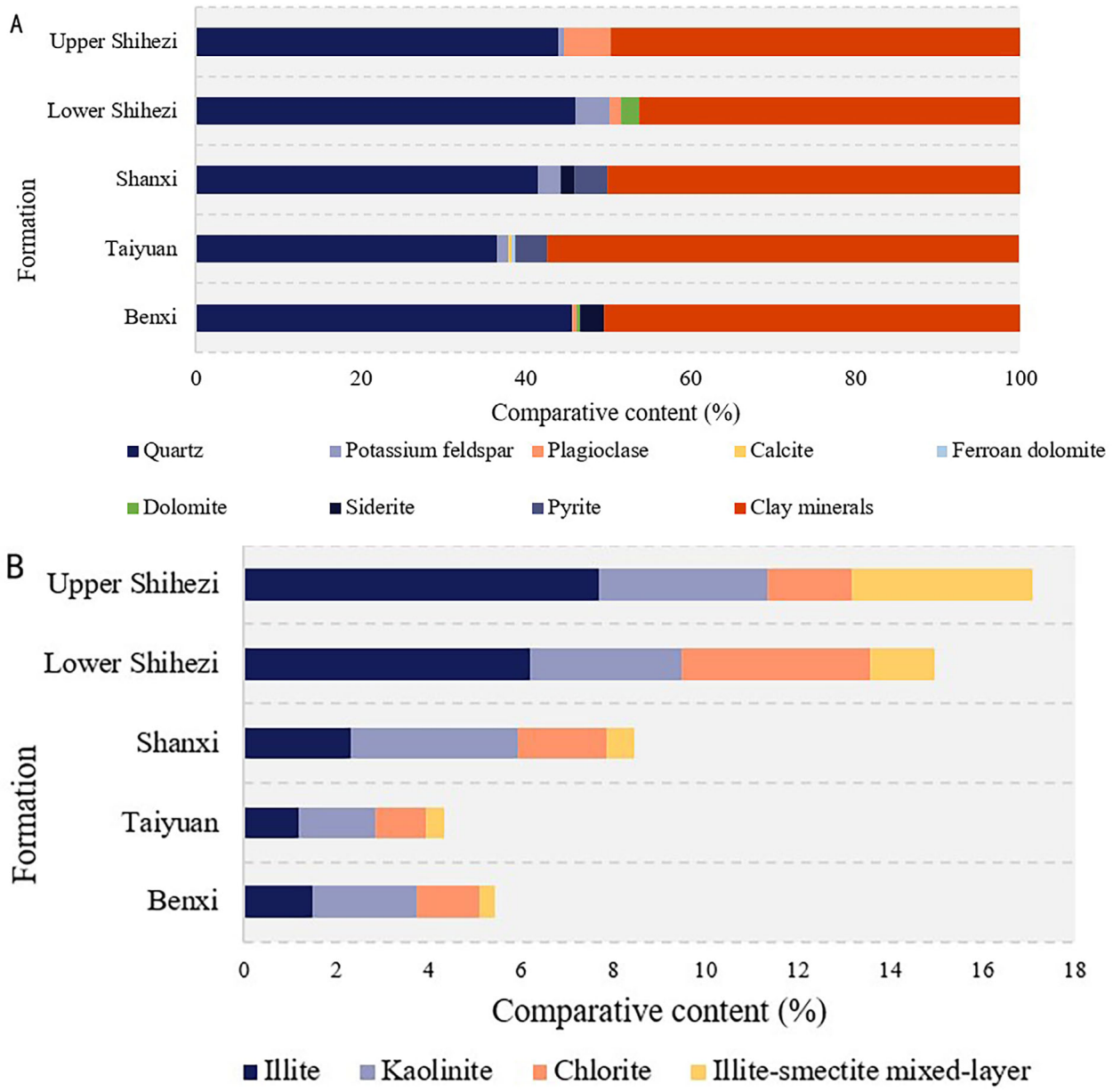



**Figure 3.** Mineral Composition of the Upper Shihezi Formation-Benxi Formation. (A) Bulk mineral composition of the samples. (B) Relative distribution of individual clay minerals within the total clay fraction.

log motifs. In the barrier island Taiyuan Formation, lagoon and swamp facies (containing coal and mudstone) corresponded to box-shaped and finger-shaped patterns, whereas tidal channel sandstones predominantly yielded a box-shaped signature. The carbonate mound facies, represented by thick marine bioclastic limestone, serves as a crucial indicator of a shallow epicontinental sea environment (Figure 4).

(ii) Facies correlation: Stratigraphic correlation based on wells LX-37, LX-34, LX-5, and LX-48 (Figures 5 and 6) indicated that the thickness of the Shanxi Formation was generally consistent. However, there was a local increase at well LX-5, which could be attributed to the development of deltaic lobes (comprising distributary channels and interdistributary bays). The Taiyuan Formation exhibits a thinning trend toward the center and thickening toward the margins, with coal seams being preferentially preserved in lagoon and swamp microfacies.

Analysis of sandbody orientation using image logs

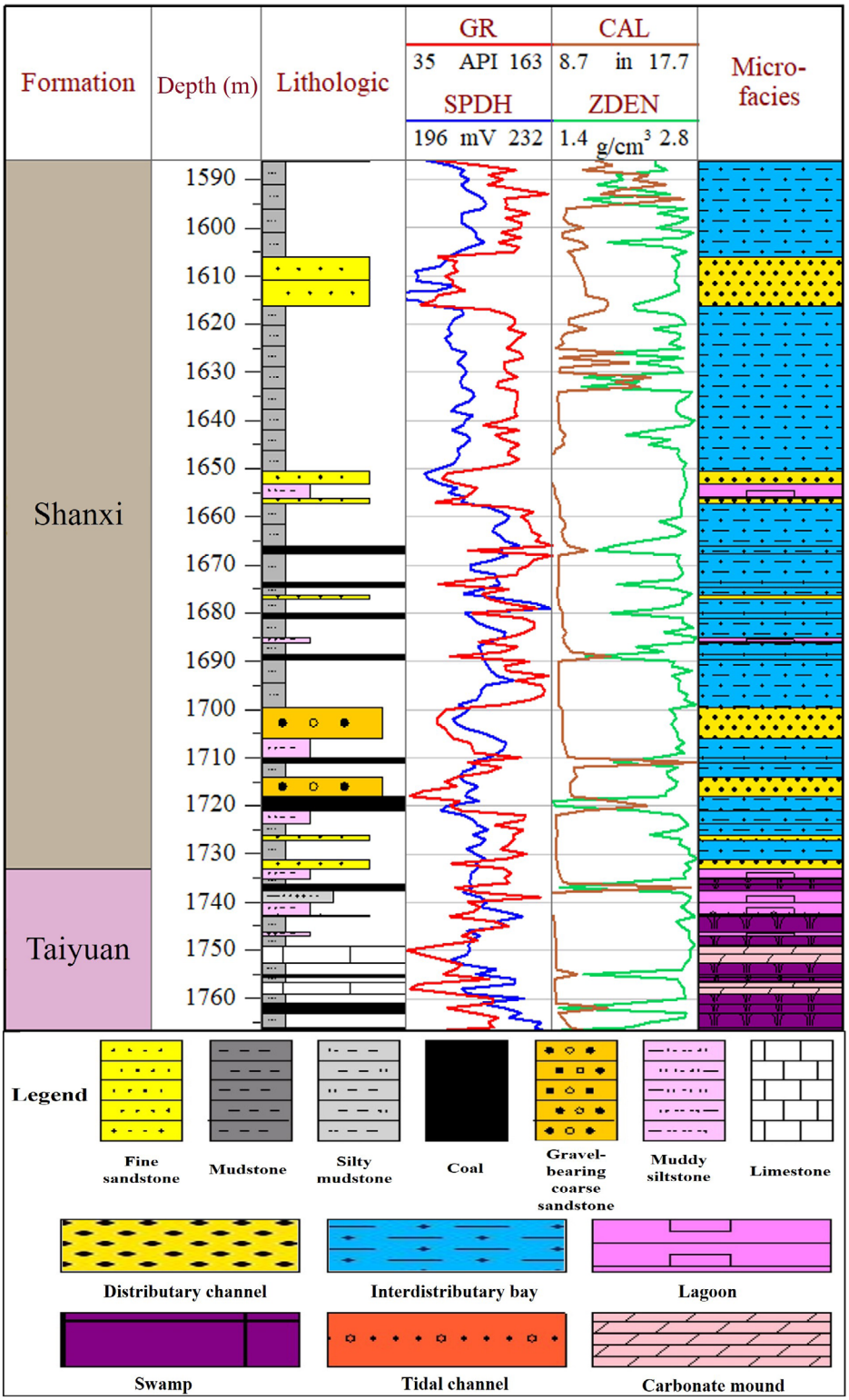



**Figure 4.** Comprehensive sedimentary facies map of the Shanxi Formation and Taiyuan Formation in well LX-5
Abbreviations: CAL: Caliper log; GR: Gamma-ray log; SPDH: Spontaneous potential from downhole tool with reference to a common remote electrode; ZDEN: Z-axis density log.

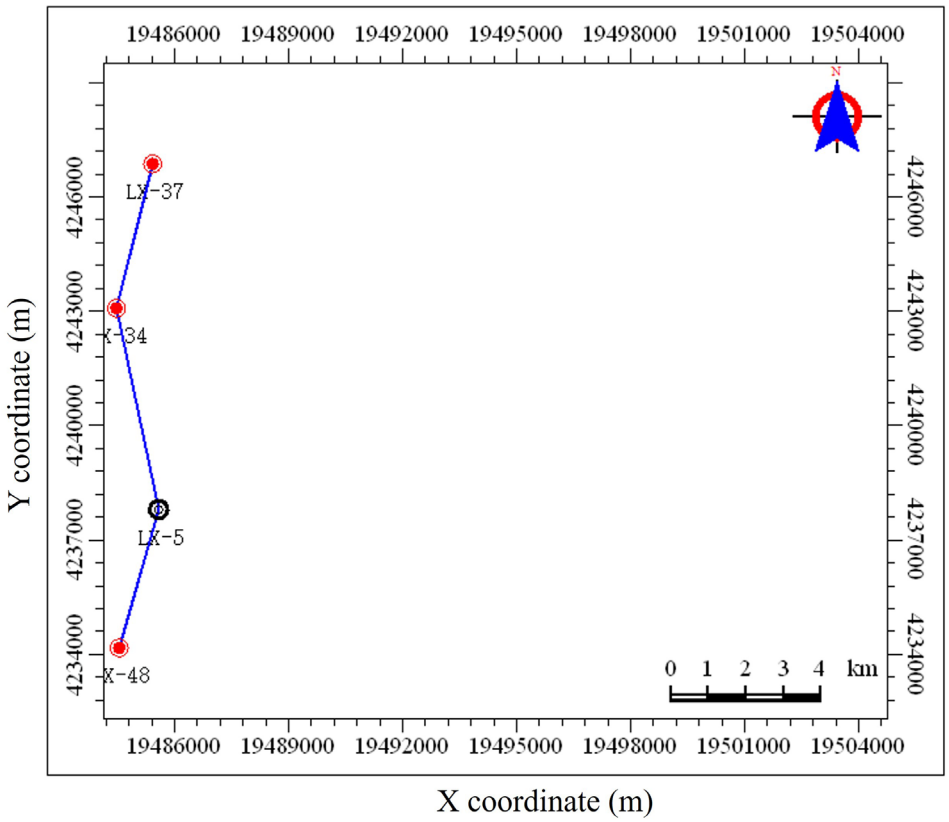



**Figure 5.** Location map of the well profiles from LX-37 to LX-48

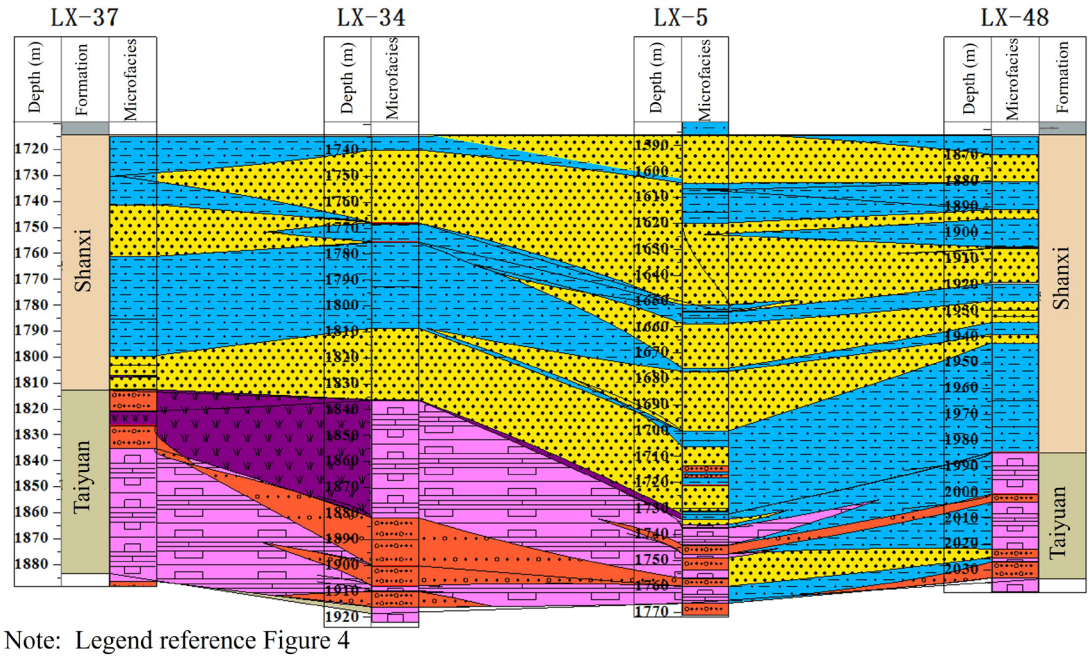



**Figure 6.** Sedimentary connected well sectional diagram of the LX-37 to LX-48 section

revealed that the primary paleocurrent direction of the Shanxi Formation was southward (locally northward), which confirmed a tide-dominated deltaic depositional environment with fluvial influence. The co-occurrence of high-energy sedimentary structures and significant coal seam thickness in the Taiyuan Formation collectively suggests deposition as a high-energy barrier island complex.

(iii) Planar facies distribution: Different seismic attributes have different sensitivities to sedimentary facies and lithology, and their prediction ability depends on the specific geological and geophysical characteristics of the study area. To objectively optimize the most sensitive attributes for the sedimentary microfacies of coal measure strata in this area, we first extracted six common attributes that may reflect lithology, physical properties, and structural characteristics based on geological knowledge: instantaneous amplitude, instantaneous frequency, instantaneous phase, sweetness, curvature, and root mean square (RMS) amplitude.

The optimization was based on two main aspects. First, from a geological–geophysical perspective, RMS amplitude reflects reflection intensity and is sensitive to the sandstone (high impedance)–mudstone (low impedance) interface, making it suitable for characterizing sandbody distributions, such as distributary channels and sand bars. The sweetness attribute, which combines amplitude and frequency, is often associated with high-porosity sandstone. Instantaneous phase highlights formation continuity and aids in tracing thin layers and lithologic boundaries. Instantaneous frequency is commonly related to formation absorption, attenuation, and tuning effects; however, in the strongly reflective coal-measure strata of this area, the strong absorption of high-frequency energy by coal seams may reduce the frequency information's resolution to lithological changes, resulting in limited predictive ability. Second, a well-seismic quantitative correlation analysis was performed. We calculated the statistical correlation between seismic attribute values and known sedimentary microfacies at key wells (LX-5). In this analysis, each depth point along the wellbore was assigned a microfacies label (categorical variable), and the corresponding attribute value at that depth was used for point-by-point correlation computation. The results show that RMS amplitude, sweetness, and instantaneous phase exhibited the strongest correlation with sandstone development degree and microfacies type, with correlation coefficients $r > 0.7$. In contrast, the curvature attribute exhibited a weak, unstable relationship with lithology/microfacies ($r < 0.3$), confirming its limited predictive capability in this dataset.

The threshold for the correlation coefficient was not a universal fixed standard, but an operational screening standard adopted for this study under specific dataset conditions. This empirical partition ($r > 0.7$ denoting "strongly correlated") highlighted the most predictive attributes, whereas the low correlation of curvature ($r < 0.3$) reflected its dataset-specific limitation. The threshold was therefore used primarily for attribute selection in this context, rather than as a general benchmark.

Consequently, based on the aforementioned geophysical principles and quantitative analysis, we optimized RMS amplitude, sweetness, and instantaneous phase attributes for subsequent attribute fusion and sedimentary facies modeling (Figure 7). The interpreted facies distribution revealed systematic seismic–geologic correlations: high RMS amplitude (red) was associated with distributary channels rich in sandstone; low amplitude (blue) corresponded to interdistributary bays filled with mud. In the Taiyuan Formation, distinct amplitude patterns distinguished the barrier island facies (high amplitude), tidal channel facies (medium-high amplitude), and lagoon facies (low-medium amplitude), as evidenced by the direct comparison between amplitude maps (Figure 7A,C) and facies maps (Figure 7B,D). The characteristic signature of the barrier island was ascribed to a strong impedance contrast at the sandstone-mudstone interface. These relationships emphasize the fundamental influence of acoustic impedance contrast on seismic amplitude in this context. Specifically, pure sandstone bodies produce high amplitudes, whereas clay-rich units produce low amplitudes.

#### *4.2.2. Sedimentation environment evolutionary process*

A multi-phase evolution of depositional environments is identified in the Upper Paleozoic sequence (Figure 8), reflecting changes in tectonic regime, paleogeography, and sediment source. (i) Late Carboniferous: Widespread marine transgression occurred, accompanied by the development of a barrier-coast lagoon system and the expansion of underwater distributary channels. (ii) Early Permian: Regional regression took place, with the development of barrier islands. The sedimentary environment was mainly dominated by the delta front. (iii) Middle Permian: The northern provenance weakened, and the sedimentary environment transitioned from a deltaic to a terrestrial lacustrine one, with the lake system gradually becoming the dominant feature. (iv) Late Permian: The sedimentary environment was dominated by terrestrial lake settings, with the extensive development of

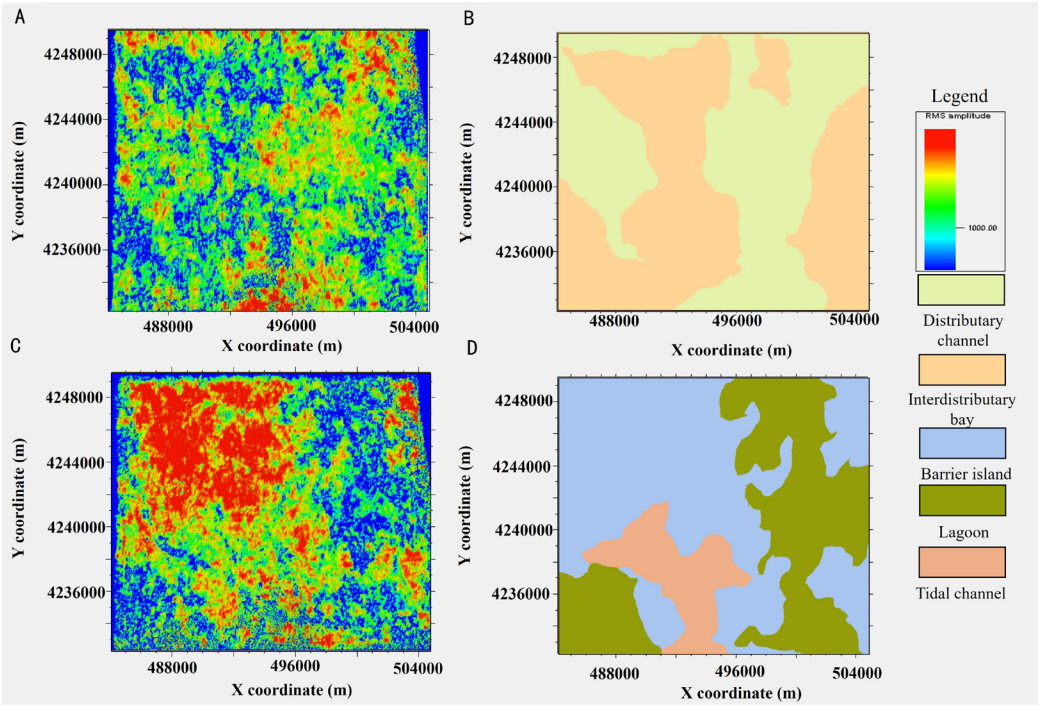



**Figure 7.** Root mean square (RMS) amplitude and sedimentary facies of the Shanxi Formation and the Taiyuan Formation. (A) RMS amplitude of the Shanxi formation. (B) Sedimentary facies of the Shanxi formation. (C) RMS amplitude of the Taiyuan formation. (D) Sedimentary facies of the Taiyuan formation.

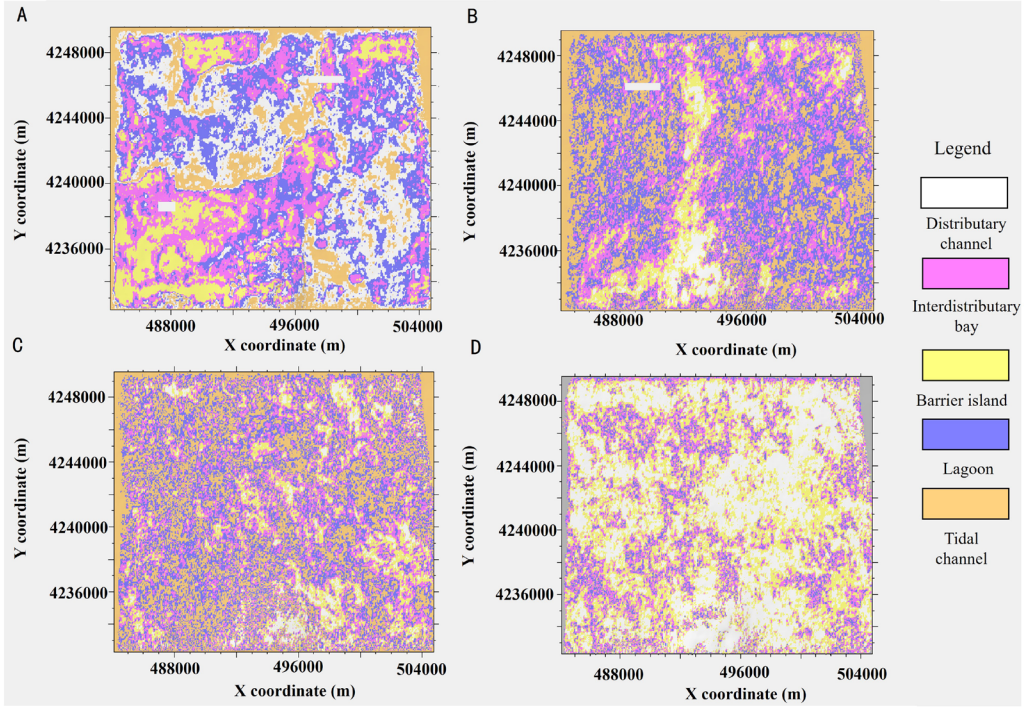



**Figure 8.** Evolution of the sedimentary environment. (A) Late Carboniferous epoch. (B) early Permian epoch. (C) Middle Permian epoch. (D) Late Permian epoch.

river deltas and shallow lakes.

#### 4.2.3. The influence of the sedimentary environment on the differentiation of metal elements (Taiyuan Formation)

The enrichment level of elements was evaluated using the concentration coefficient (CC), defined as the ratio of an element's content in the studied coals to its average content in world coals.[36] According to the CC values, elements can be classified into four categories: enrichment ($5 < CC \leq 10$), slight enrichment ($2 < CC \leq 5$), normal ($0.5 < CC \leq 2$), and loss ($CC \leq 0.5$) (Figure 9).

The application of this scheme to the average metal content in the Shiqianfeng Formation–Benxi Formation of well LX-5 indicates that Zr exhibits enrichment; Cr, nickel (Ni), and hafnium (Hf) display slight enrichment; Li, scandium, vanadium (V), cobalt (Co), copper (Cu), zinc, Ga, rubidium (Rb), Sr, niobium, indium, cesium, Ba, and U are at background levels; while beryllium and cadmium show depletion.

For the study area's coal measures, paleoenvironmental conditions were reconstructed using combined petrographic, geophysical, and geochemical analyses. This discussion leverages that reconstruction to examine how sedimentary evolution influenced metal distribution. Table 2 presents the calculated trace element data for some LX-5 samples.

(i) Palaeosalinity (Sr/Ba): Sr/Ba ratios derived from key formations suggest a freshwater paleosalinity regime in the Taiyuan Formation (Figure 10). This geochemical evidence validates the interpreted tidal delta-barrier island environment, which underwent episodic marine incursions within a broader epicontinental sea context. The consequent extensive development of tidal and lagoon facies under such fluctuating salinity conditions fostered a reducing environment, thus facilitating the preservation of organic matter.

(ii) Palaeoclimate: The Sr/Cu ratio is a sensitive paleoclimatic indicator, reflecting the balance between elements for dry and wet conditions. Sr/Cu < 10 means a warm and humid climate, while Sr/Cu > 10 indicates a dry and hot climate.[37] The distribution of Sr/Cu values (Figure 11) shows that the Taiyuan Formation's sedimentary period was mainly warm and humid, consistent with it being a tidal delta barrier coastal sedimentary system in a land-surface sea environment. Additionally, variations in high and low S content of samples were investigated (Table 1). The Taiyuan Formation was mainly affected by terrestrial inputs and exhibited intermittent marine transgression features. Geochemical and sedimentary data together showed that paleoclimatic evolution, driven by sea-level fluctuations and changes in source supply, fundamentally controlled the geochemical signature of the coal measures.

(iii) Oxidation-reduction conditions: The distribution of redox-sensitive trace elements (e.g., V, Ni, Cr, U) in the samples is predominantly regulated by depositional redox conditions, as corroborated by their systematic variations across facies.[38,39] Table 3 presents an index for evaluating the redox environment of trace elements.

The V/Cr and Ni/Co ratios suggest that the overwhelming majority of samples were situated within an oxidizing environment. Merely a minuscule quantity of samples (e.g., L5-73) were in a weakly oxidizing-weakly reducing environment. In contrast, the L5-18 samples exhibited a reducing environment due to hydrothermal activity. A comprehensive analysis demonstrated that the sedimentary period in the study area was predominantly characterized by oxidative conditions, with a scarcity of typical and extensive reducing environments.

#### 4.2.4. Explanation of the reduction signal contradiction

In this study, redox indicators showed a complex situation. Core observations of dark lithology, high TOC content, and widespread pyrite suggest a reducing depositional environment.[40,41], with geochemical ratios (e.g., V/Cr and Ni/Co) mostly indicating oxidation. This discrepancy implied that subsequent geological processes may have overprinted original sedimentary signals. The following possibilities were explored to explain these inconsistencies[39]:

(i) Yanshanian hydrothermal overprinting: Extensive Yanshanian tectonic-thermal events in North China might have altered original geochemical signals. Oxidizing and acidic hydrothermal fluids could have dissolved primary reducing minerals and introduced or modified elements, resetting V/Cr and Ni/Co ratios to suggest oxidative conditions.[42]

(ii) Terrestrial detrital dilution: The study area on the basin margin was affected by terrestrial input. Detrital minerals carrying Cr and Co diluted the authigenic signal of redox-sensitive elements, resulting in lower ratios and a false indication of an oxidative environment.[42]

(iii) Proxy limitations in multi-stage basins: Classic redox proxies like V/Cr and Ni/Co, established for marine shale, required careful assessment when applied to inland basins with complex diagenesis and multi-stage tectonic events. Multiple indicators should be used for verification in multi-cyclic basins.[42]

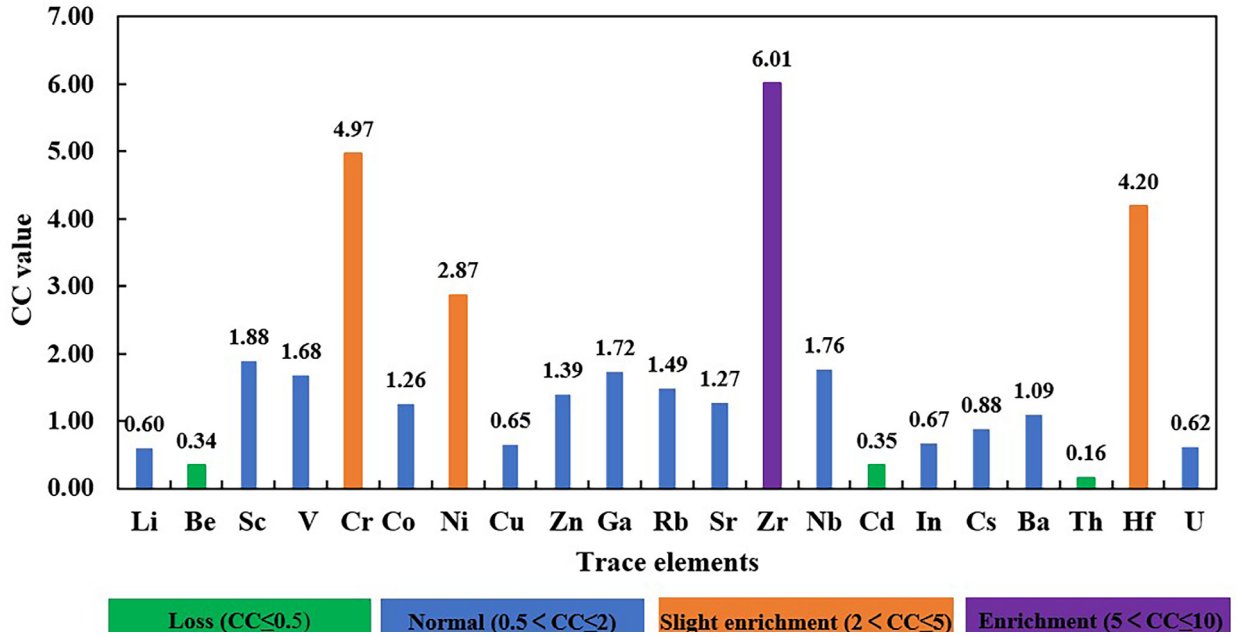



**Figure 9.** The calculated trace element concentration coefficient (CC) of the LX-5 well

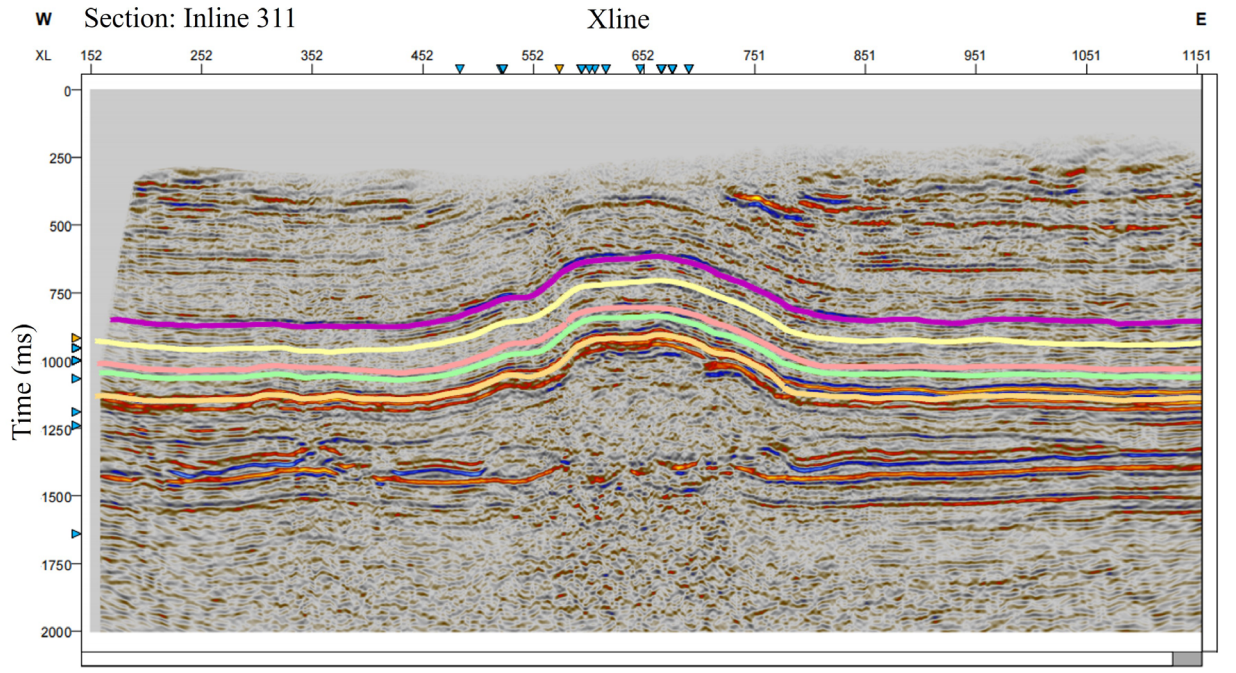



**Figure 10.** Seismic interpretation layer. Strata from the top to the bottom are: Shiqianfeng (P3s), Upper Shihezi (P2s), Lower Shihezi (P2x), Shanxi (P1s), and Benxi (C2b) formations.

**Table 2. Trace element analysis calculation (μg/g)**

| Sample | Formation | Sr/Ba | Sr/Cu | V/Cr | Ni/Co | V/(V + Ni) | Co | Th/U |
|---|---|---|---|---|---|---|---|---|
| L5-7 | Shiqianfeng | 0.94 | 46.50 | 1.90 | 1.70 | 0.84 | 3.20 | 1.15 |
| L5-11 | Upper Shihezi | 2.25 | 126.10 | 0.44 | 4.46 | 0.50 | 5.06 | 3.93 |
| L5-18 | Lower Shihezi | 0.14 | 1.30 | 0.10 | 16.60 | 0.14 | 11.05 | 3.52 |
| L5-22 | Taiyuan | 0.34 | 16.30 | 1.48 | 2.19 | 0.71 | 3.66 | 3.12 |
| L5-73 | Benxi | 0.63 | 2.50 | 3.78 | 1.67 | 0.84 | 14.71 | 5.15 |

Abbreviation: Ba: Barium; Co: Cobalt; Cu: Copper; Ni: Nickel; Sr: Strontium; Th: Thorium; U: Uranium; V: Vanadium.

**Table 3. Trace elements redox environment judgment index**

| Trace element | Reduction | Weak oxidation–weak reduction | Oxidation |
|---|---|---|---|
| V/Cr | >4.25 | 2.00–4.25 | <2 |
| Ni/Co | >7.00 | 5.00–7.00 | <5 |

Abbreviations: Co: Cobalt; Cr: Chromium; Ni: Nickel; V: Vanadium.

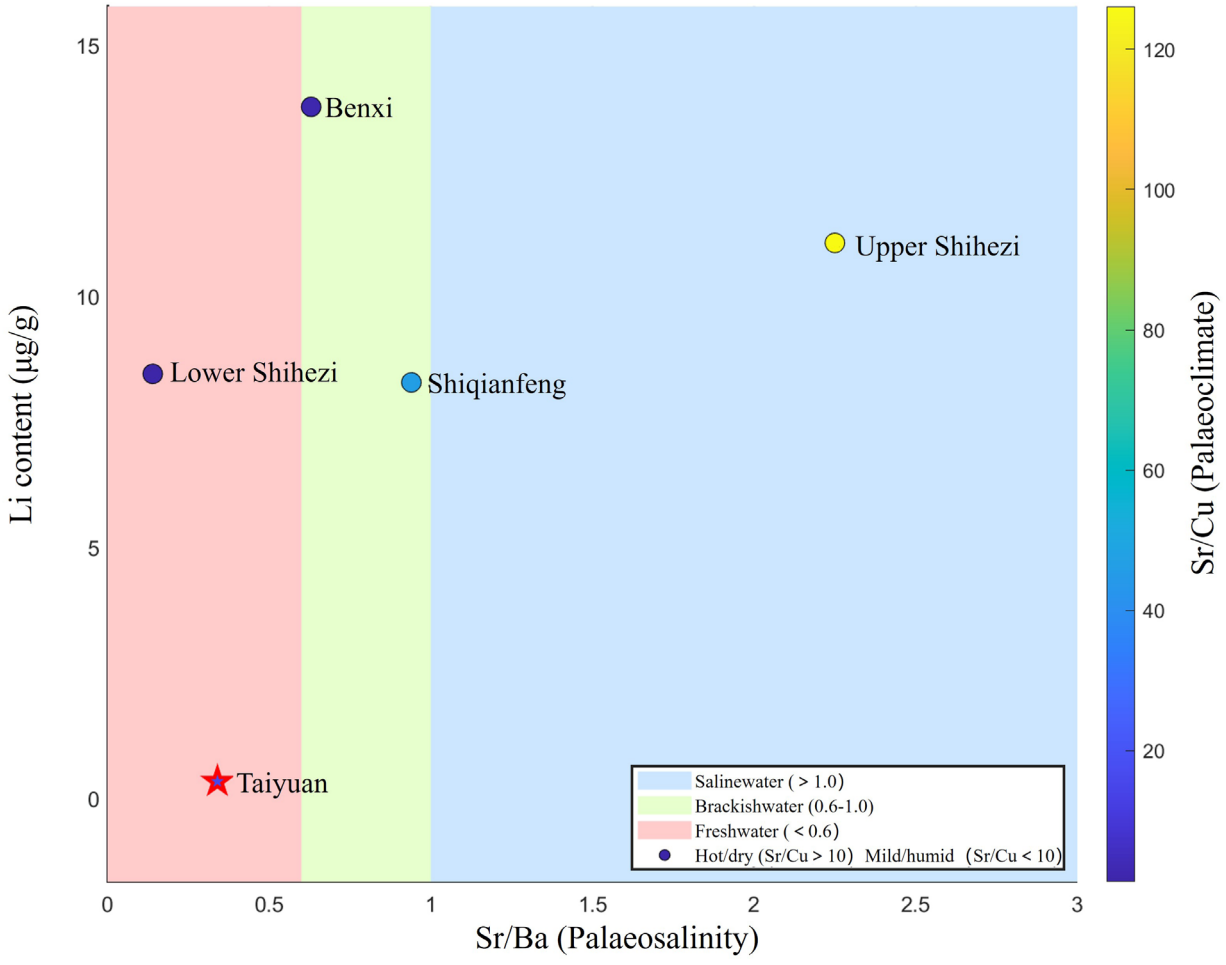



**Figure 11.** The distribution of lithium (Li), strontium (Sr)/ barium (Ba), and Sr/ copper (Cu)

Based on the evidence, the Taiyuan Formation coal seams formed in an anoxic environment during sedimentation and early diagenesis, which is favorable for organic matter preservation and pyrite formation. However, subsequent hydrothermal activity (likely Yanshanian) modified the original geochemical features. Thus, primary sedimentary environment information from core observations and TOC is more reliable than potentially modified elemental ratios.

#### *4.2.5. Multi-factor control on metal enrichment*

Integrated analyses demonstrated that the sedimentary environment of the Taiyuan Formation had multi-factor control on metal differentiation. Facies analysis indicated that the Taiyuan Formation developed diverse facies, including lagoon and tidal-flat settings, whereas the overlying Shanxi Formation had a homogeneous, river-dominated delta system. Paleosalinity and lithofacies, combining with trace-element indicators, suggest that the Taiyuan Formation was deposited in a tidal delta-barrier island system in an epicontinental sea. This freshwater-dominated system, interrupted by transgressions, enabled the preservation of organic matter and the development of reducing conditions. The paleoclimate was warm and humid, and together with high S content and peat swamps, it enhanced the stability of the reducing interface. The reducing environment of the Taiyuan Formation regulated the evolution of organic matter and influenced metal enrichment through sulfidation and organometallic complexation.

### 4.3. Tectonic thermal events

#### *4.3.1. Balanced cross-section*

The construction of balanced cross-sections depends on quantitative erosion (acoustic time-difference method) and on reconstructing the burial history. Integrating interpreted seismic horizons (from the Heshanggou Formation to the Benxi Formation) (Figure 10) and well-based depth conversion indicates a structure modified by uplift and erosion. Therefore, a balanced cross-section (Figure 12) was constructed by restoring the missing section, decompacting strata, and sequentially removing fault displacement and folding.

The tectonic history of the study area was defined by a multi-phase evolution with major unconformities (Figure 12). It began with the deposition of Ordovician carbonate platform sediments in the early Paleozoic. The Caledonian Orogeny ended this phase, causing uplift and significant erosion that removed Silurian and Devonian strata. Then, in the late Paleozoic, Variscan-influenced subsidence allowed the deposition of Carboniferous–Permian coal measures and overlying Permian–Triassic clastic rocks on the truncated Ordovician basement. Triassic burial compacted the coals rapidly under stable conditions. In the Mesozoic, a major tectonic inversion occurred. The Indosinian and Yanshanian orogenies led to the early Cretaceous uplift of the Zijinshan belt, compression, reactivation of fault systems, and erosion of Jurassic–Cretaceous deposits. A weak extensional regime then stabilized the structure. Finally, from the Cenozoic to the Quaternary, there was tectonic quiescence, with thin Tertiary and Quaternary sediments covering the eroded Triassic landscape.

#### *4.3.2. Zijinshan pluton*

The Late Jurassic to Early Cretaceous period was characterized by intense magmatism, as exemplified by the Zijinshan pluton in the southern study area. This intermediate-to-deep intrusive-extrusive complex intruded the Lower Paleozoic Majiagou Formation. Remnants of carbonate roof pendants and overlying Carboniferous–Permian strata showed their post-Paleozoic emplacement (Figure 13).

This magmatic activity affected metal endowment through a multi-faceted process. A high geothermal gradient around the Zijinshan pluton indicated a thermal anomaly, which accelerated organic matter maturation in adjacent coal measures and raised vitrinite reflectance in the Benxi Formation to 0.978–1.654% (Table 1). Then, magmatically derived hydrothermal fluids circulated through the sequence, transporting, introducing, and enriching (or depleting) trace metals in coal seams.[43–46]

Moreover, the emplacement dynamics created abnormal fluid pressures, forming and reactivating fracture networks in multiple stratigraphic units. Later-stage volcanic eruptions increased permeability by dilating fractures and creating secondary cracks. These structures were the main conduits through which metal-laden hydrothermal fluids entered coal-bearing strata, especially along fault zones and volcanic conduits.

#### *4.3.3. Fluid inclusion constraints on hydrothermal activity*

Microthermometric analysis of fluid inclusions from quartz grains in the Shanxi Formation sandstones of well LX-5 showed direct constraints on the nature and timing of hydrothermal activity. Two populations of secondary aqueous inclusions were identified, hosted within healed micro-fractures and dissolution pores. They recorded distinct thermal events: an earlier, lower-temperature phase (Event I: Th 92.8–185.7 °C) and a later, dominant moderate-temperature phase (Event II: Th 104.7–183.2 °C), both exhibiting moderate to high salinities (6.74–12.96

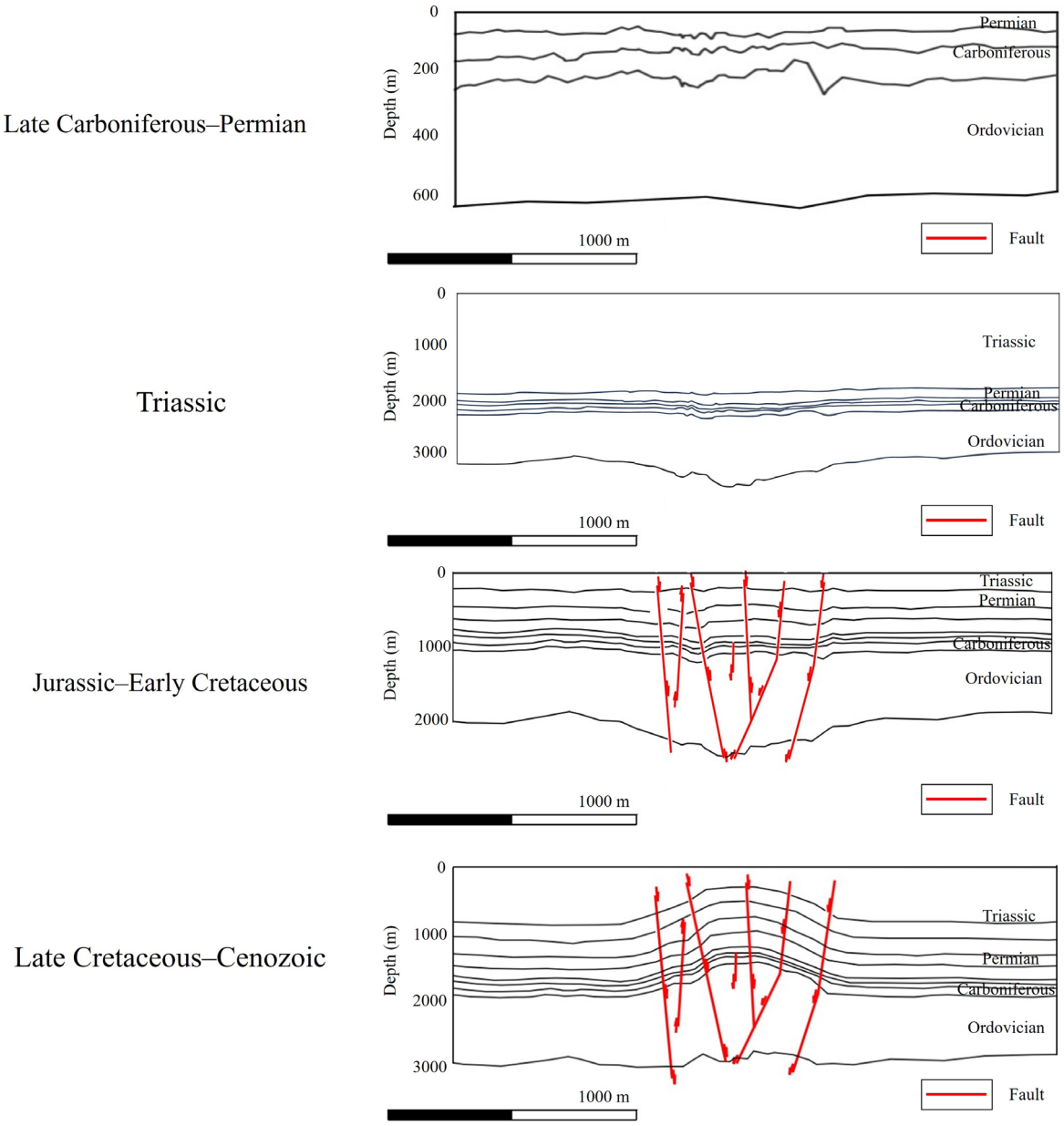



**Figure 12.** Tectonic evolution process in the study area based on the balanced cross-section technique

weight % sodium chloride equivalent). The temperature range of Event II aligned with the moderate-temperature hydrothermal regime independently inferred from elevated vitrinite reflectance and REE signatures. Critically, the secondary nature of these inclusions, their textural association with late-stage fractures, and their moderate-temperature–moderate-salinity characteristics collectively indicated that the entrapped fluids were genetically linked to the Yanshanian tectono-thermal event rather than to earlier diagenetic or later unrelated fluid episodes. This evidence reinforces that fault and fracture networks acted as primary conduits for metal-bearing hydrothermal fluids during the Yanshanian, thereby explaining the post-depositional remobilization and localized enrichment of critical metals such as Cr.

#### *4.3.4. Mechanisms and controls of critical metal enrichment in coal strata*

The enrichment of critical metals in coal strata is attributed to the coupled effects of recurrent magmatism and exhumation. This coupling occurs through distinct but complementary processes: (i) Syn-depositional fault

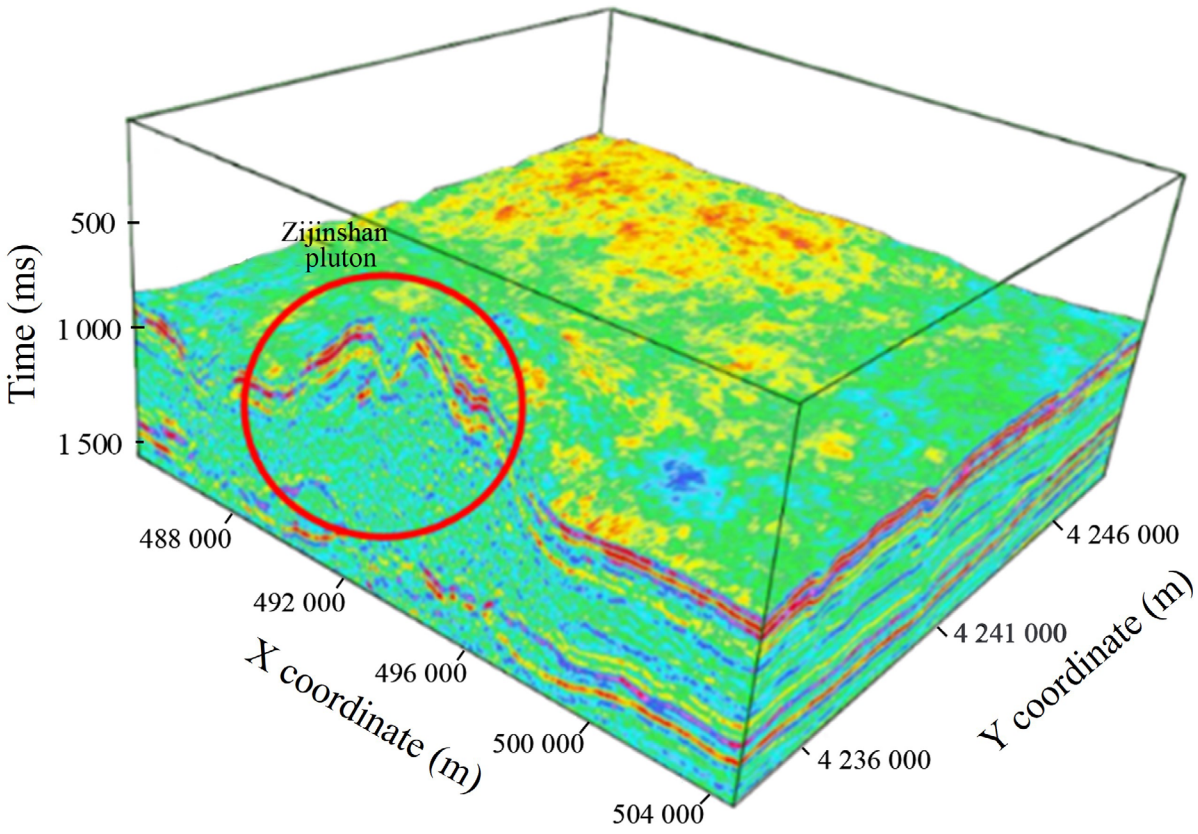



**Figure 13.** Three-dimensional visualization of Zijinshan pluton interpretation

**Table 4. Summary of fluid inclusion data by event**

| Event | Number of inclusions | Homogenization temperature range (°C) | Mean temperature (°C) | Salinity range (weight %) | Mean salinity (weight %) | Main occurrence | Host mineral |
|---|---|---|---|---|---|---|---|
| Event I | 10 | 92.8–185.7 | 131.67 | 8.41–10.61 | 9.56 | Fissure | Quartz |
| Event II | 10 | 104.7–183.2 | 145.22 | 6.74–12.96 | 9.94 | Fissure, dissolution | Quartz |

networks, activated by the Zijinshan pluton, promoted topographic differentiation and concentrated the vertical migration of deep ore-bearing fluids, resulting in metal deposition at permeability barrier islands within the coal sequence; (ii) The direct influx of hydrothermal fluids containing Au, Cu, and REEs facilitated alteration assemblages (e.g., kaolinite), as evidenced by elemental anomaly studies.

The significance of integrating multiple controlling factors is exemplified by the LX-5 well, situated in a tectonically stable area far from volcanic centers. Its distinct trace element and REE patterns, in contrast to those in the volcanically influenced zone, cannot be explained solely by the structural history. This geochemical disparity conclusively indicates that an effective prediction of metal-enrichment heterogeneity requires a model that simultaneously accounts for sedimentary facies distributions and the influence of paleo-groundwater systems.

### 4.4. Groundwater distribution

The distribution and flow of groundwater, which are regulated by stratigraphic porosity and permeability, tectonic fractures, and hydrological gradients, served as the primary medium for the activation, migration, and reprecipitation of critical metals. The subsequent analyses reconstructed the paleogroundwater system and delineated its specific roles in metal enrichment.

#### *4.4.1. Wave impedance inversion*

This study used wave impedance inversion technology to identify high-porosity and high-permeability sandstone bodies (e.g., distributary channels and sand dams) by precisely characterizing subsurface lithology wave impedance differences. Such sandstone bodies are favorable for controlling critical metal migration and enrichment in coal measures; thus, the inversion results directly contribute to predicting the spatial distribution of critical metals.

Traditional inversion methods are limited by wavelet and model constraints.[47,48] Recent deep learning approaches, particularly generative adversarial networks (GANs), enhance automation and resolution by learning nonlinear seismic-impedance mappings.[49] Unsupervised physics-driven networks[50] and GAN-based data augmentation methods[51] further enhance the robustness of the methods under limited samples.

The wave impedance value was obtained by natural gamma pseudo-acoustic inversion. The black-lined low-value green area was sandstone, and the rest was mudstone. The low-impedance zone (green) in Figure 14 was highly consistent with the sandstone development zone interpreted from logging, confirming its reliability. These confined low-impedance sandstone bodies, with high porosity and permeability, are interpreted as fluid-dominant channels for the migration and enrichment of metal elements, including Cr, Ni, and Zr, providing key geophysical evidence for constructing a tectonic-sedimentary fluid-joint ore-control model.

#### *4.4.2. Property–structural coupled model*

Model construction and integration used geostatistical simulation of log data to construct 3D porosity-permeability models, which were overlaid on structural maps (Figures 15 and 16). Comparative analysis with the sedimentary facies model (Figure 7) indicated that high-quality reservoir zones corresponded spatially with distributary channel and tidal bar sand bodies, suggesting depositional facies controlled primary pore development. Structural elements, especially faults in the Zijinshan pluton, further conditioned these properties. High-porosity regions often correlate with structural highs and are inferred as groundwater conduits.[52] Fault-mediated enhancement varies by lithology, improving sandstone porosity and fracturing muddier units. This synthesis of sedimentological and structural controls forms the basis for a two-factor reservoir evaluation model.

The property–structural coupled model is relevant to critical metal enrichment as porosity and permeability control hydrothermal fluid flow in the coal-bearing sequence. High-porosity and high-permeability zones (e.g., distributary channel sand bodies) are preferential conduits for metal-bearing fluids, while low-permeability mudstone layers can act as seals. Faults and fracture zones within the structural coupling are critical fluid pathways that connect deep-seated metal sources to reservoir rocks. Thus, the coupled spatial distribution of porosity, permeability, and the structural framework provides a predictive model for critical metal sweetness, which underpins the interpretation of observed metal distribution patterns.

#### *4.4.3. Groundwater leaching and weathering*

An anomalously high Th/U ratio (>7) is typically due to groundwater leaching under strong oxidation. However, hydrothermal fluids can also leach and transport U, leading to U loss and increasing the Th/U ratio. The Co content can be used for auxiliary discrimination, as hydrothermal activity can lead to significant Co enrichment. The Th/U ratios of all samples were calculated (Figure 17) to identify samples with anomalously high (>7) or low (<2) values. The H-35 sample of Upper Paleozoic mudstone had a Th/U value of 39.43, indicating strong oxidation and groundwater leaching.

These data showed that source-area rocks generally weather highly, and leaching by atmospheric freshwater could enrich and differentiate REEs in source-area granite. By combining the Th/U ratio and Co trace-element data, a comprehensive interpretation was made to clarify the combined control of redox conditions, groundwater leaching, and hydrothermal activity.

#### *4.4.4. Reconstruction of the paleo-groundwater system and its role in key metal enrichment*

This section reconstructed the paleo-groundwater system in the study area via wave impedance inversion, attribute–structure coupling modeling, and geochemical index analysis, and disclosed its core function in the activation, migration, and reprecipitation of key metals. The findings indicate that the enrichment of key metals results from sedimentation-structure-fluid coupling, deepening understanding of the mineralization of coal-measure key metals.

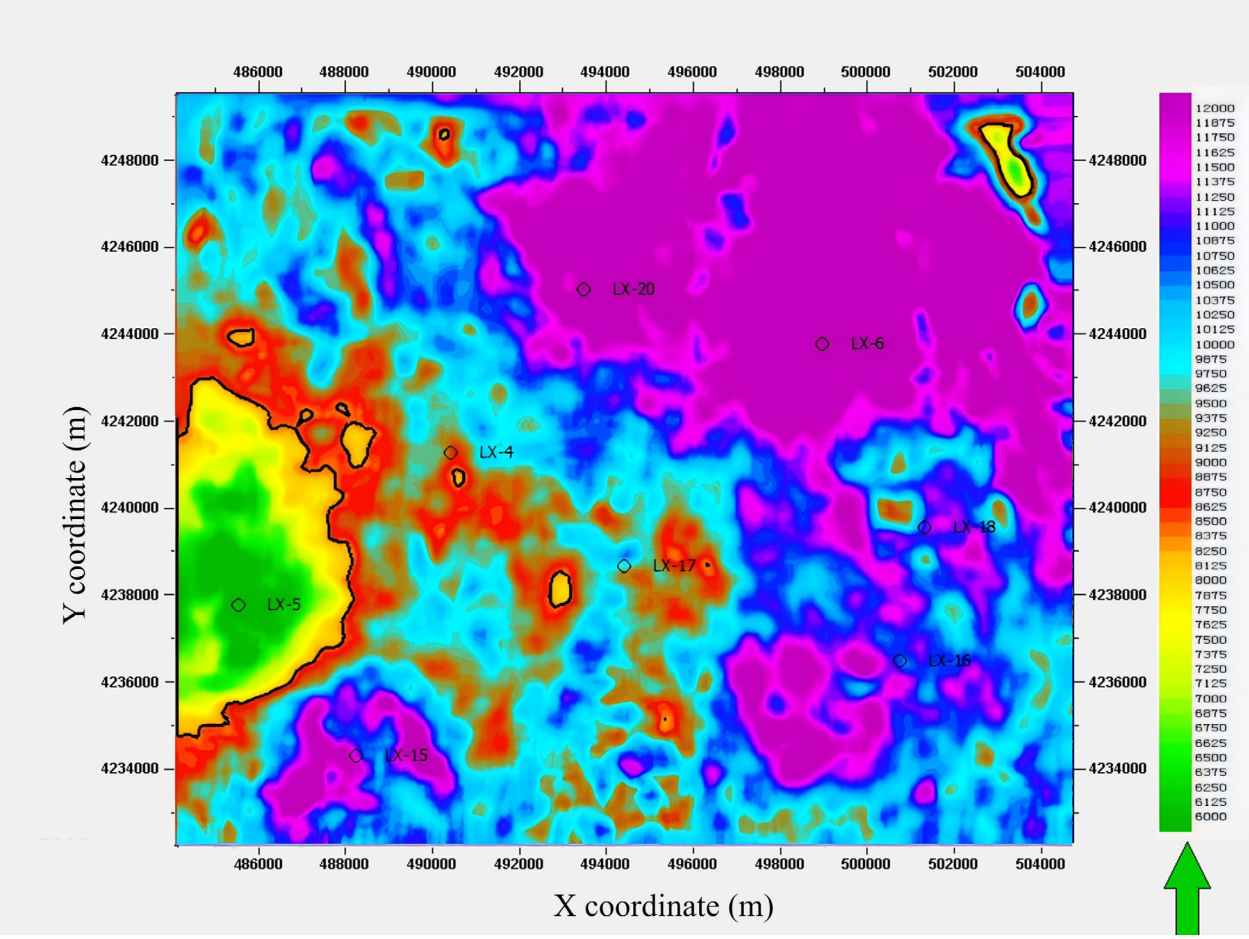



**Figure 14.** 4 + 5 # coal seam (a major coal seam within the coal-bearing strata of the study area characterized by relatively thin thickness with significant lateral variation) roof lithology

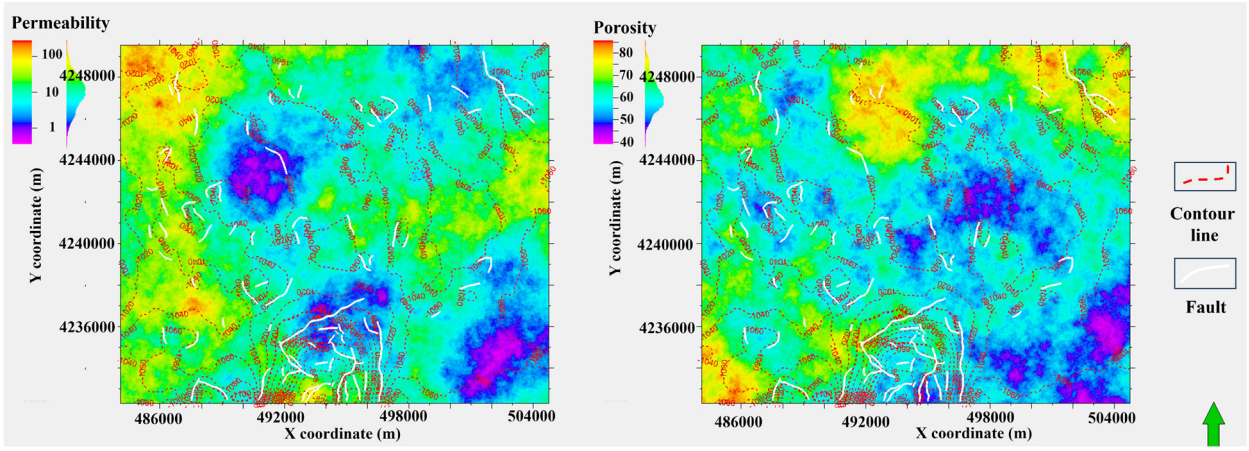



**Figure 15.** Comprehensive property–structural coupled model map of the Shanxi Formation

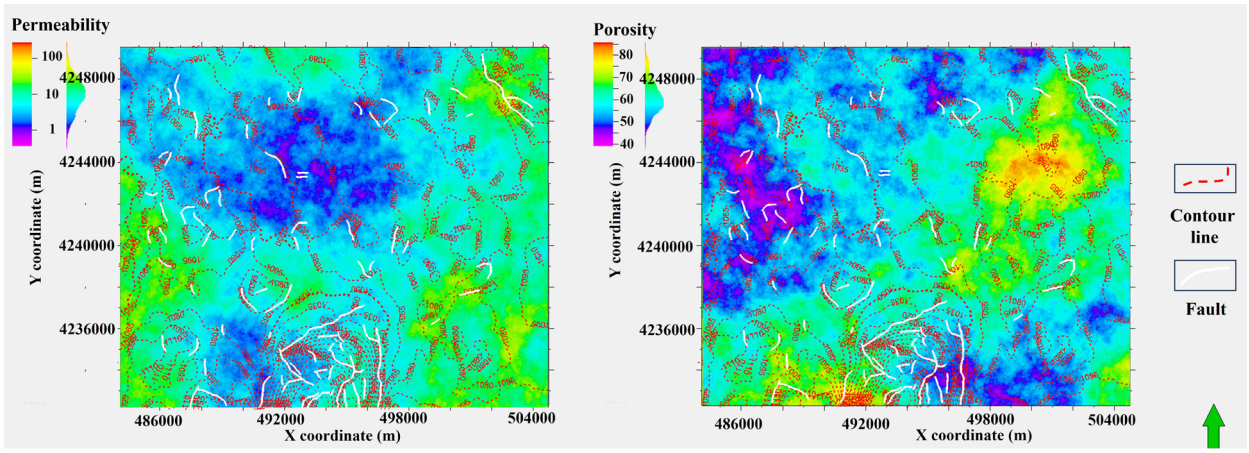



**Figure 16.** Comprehensive property–structural coupled model map of the Taiyuan Formation

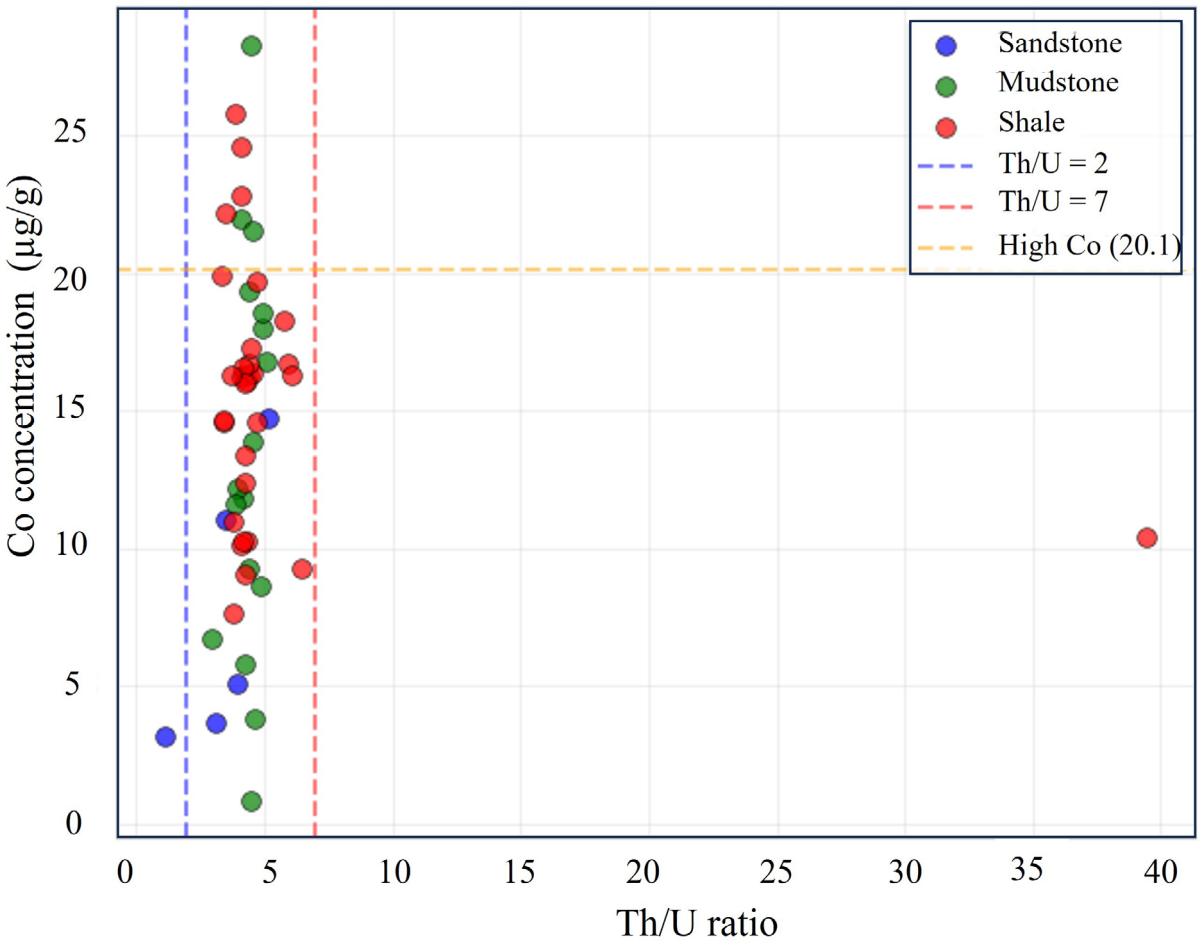



**Figure 17.** The thorium (Th)/uranium (U) ratio and cobalt (Co) content distribution of the samples

Firstly, wave impedance inversion identified sandstone bodies with high porosity and permeability, consistent with the reservoirs interpreted from well-logging data. These sandstone bodies serve as "preferential channels" for the migration of metal elements, offering evidence for the structure–sedimentation–groundwater joint ore-controlling model and establishing a connection between sedimentary pore networks and fluid transport pathways.

Secondly, the attribute–structure coupling model realized 3D quantitative characterization of the porosity-permeability field. High-quality reservoir areas were correlated with sedimentary facies, validating that sedimentary facies were the foundation for the development of primary pores. Tectonic activity modified the sedimentary framework, forming a pattern of sedimentary control and structural alteration of storage. The coupling of the porosity-permeability field and the structural framework formed the physical basis for predicting key metal "sweet-spot areas" and accounted for the distribution of metals.

Finally, geochemical tracing of groundwater leaching and weathering provided evidence of fluid-rock interactions for the physical model. An abnormal Th/U ratio indicated groundwater leaching under strongly oxidizing conditions. By analyzing trace elements such as Co, the effects of atmospheric freshwater leaching and hydrothermal activity can be discriminated. These data demonstrate that source-area rocks have experienced intense weathering, and atmospheric freshwater leaching enriches and differentiates REEs. This clarifies the chemical properties of the paleo-groundwater and integrates fluid-chemical reactions with physical transport channels.

### 4.5. Comprehensive control factors of metal distribution

The fault and fracture network generated by the Yanshanian tectonic thermal events (particularly those associated with the Zijinshan rock mass) not only functioned as a crucial conduit for hydrothermal fluids but also fundamentally remodeled the flow path of paleo-groundwater systems, significantly enhancing the stratum's permeability. These structurally controlled permeable zones directed the flow of paleo-groundwater (comprising atmospheric precipitation and deep-basin fluids) and formed and governed the position of redox fronts at the interfaces between highly permeable sandstone bodies (e.g., distributary channels) and low-permeability mudstone interlayers, or at different lithological/fluid contact zones.

The reactivation, migration, and ultimate precipitation of metals resulted from the combined effect of structural channels and porosity/permeability fields controlled by sedimentary facies. Specifically, deep hydrothermal and basin fluids migrated upward along faults (under structural control), then laterally in high-porosity sandstone bodies determined through sedimentary facies (under sedimentary control), and finally precipitated at the redox front located via structural uplift or lithological traps (under the joint control of structure, sedimentation, and fluid).

This joint-control model of structure, sedimentation, and groundwater effectively linked deep-seated hydrothermal events with shallow-fluid redistribution processes, thereby elucidating the causes of spatially heterogeneous metal enrichment.

### 4.6. Enrichment mechanism of chromium

To clarify the enrichment mechanism of critical metals in the eastern margin of Ordos Basin, Cr was selected for in-depth analysis. The reasons were as follows: First, CC indicated that Cr was slightly enriched in the study area, much higher than the background value, making it suitable for exploring metal enrichment. Second, Cr showed a specific enrichment in the Lower Shihezi Formation, providing a good opportunity to analyze geological factors, such as sedimentary environment, tectonic activity, and fluid migration. Thus, this study aims to establish a widely applicable source–transport–storage model by analyzing Cr's enrichment mechanism, applicable not only to Cr but also to other enriched critical metals such as Zr, Ni, and Hf in this area.

#### *4.6.1. Provenance analysis*

Chromium distribution within the Carboniferous–Permian succession revealed a distinct enrichment anomaly in the Lower Shihezi Formation (Figure 18).

In geochemistry, Cr is usually closely associated with ultrabasic and mafic rocks, such as basalt and serpentinite suites. Its enrichment in sediments may indicate the contribution of basic volcanic rocks, gabbro, or metamorphic basic rocks in the source area. Cr can also occur in heavy minerals, such as chromite, and provides a good source tracer. Data on Taiyuan Formation coal-bearing shale and Upper Paleozoic mudstone from previous studies were selected[12,40], and the correlation among Cr content, aluminium oxide ($Al_2O_3$), and potassium oxide ($K_2O$) in the samples was poor (Figure 19). Cr and V (with a high correlation coefficient), Ni, Co, and other siderophile elements often exhibit good positive correlation (Figure 20). The abnormally high Cr content, no significant correlation between Cr and $Al_2O_3$ and $K_2O$, and a strong correlation between Cr and V and Ni suggest that they may jointly originate from weathering products

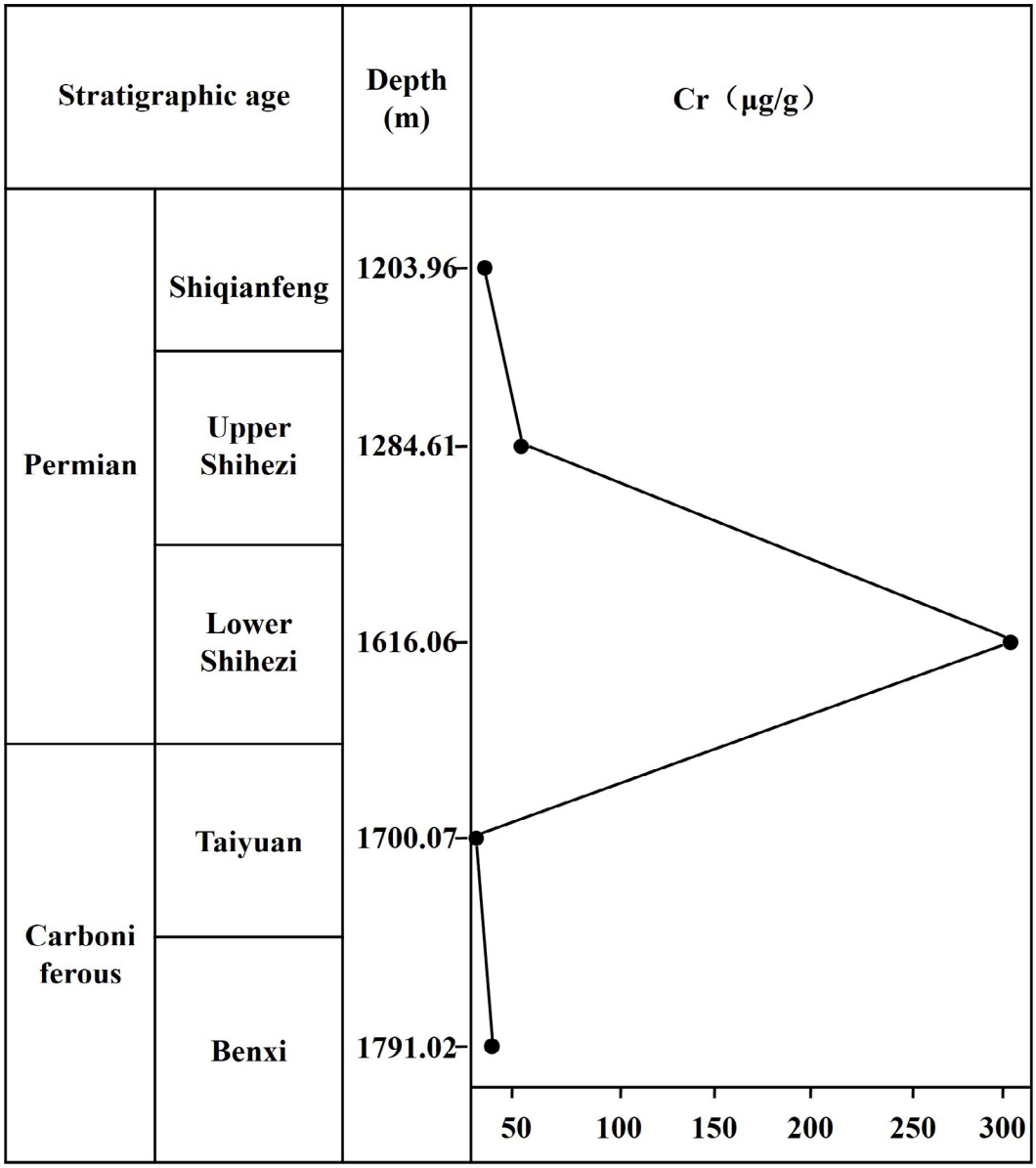



**Figure 18.** The distribution of chromium (Cr) element content

of basic/ultrabasic rocks (such as gabbro and peridotite), or be related to volcanic materials and hydrothermal activity.

The main source of Cr in coal in the study area is terrestrial debris rich in basic/ultrabasic rock components. These rocks are rich in primary minerals, such as chromite and chromite silicates. Important local sources include hydrothermal activity and volcanic ash. For the anomaly point L5-18 (sandstone), its extremely high Cr content and Cr/Th ratio, as well as the possible existence of special lithology or sedimentary structures in this layer, strongly suggest hydrothermal fluid activity during the same or quasi-same sedimentary period, or direct deposition of basic volcanic ash. Sandstone, as a highly permeable rock layer, is more likely to become a hydrothermal channel. Part of Cr may be adsorbed onto Fe oxides or bound to organic matter, but these processes are not the dominant factor.

#### *4.6.2. Distribution mode of rare earth elements*

The source of REE was determined using the anomalies of Eu, Ce, lanthanum (La), yttrium (Y), and holmium (Ho), with the formula as follows (**Equations 1–3**):

$$\delta Eu = \frac{Eu}{Eu^*} = \frac{Eu_N}{\sqrt{Sm_N \times Gd_N}} \quad (1)$$

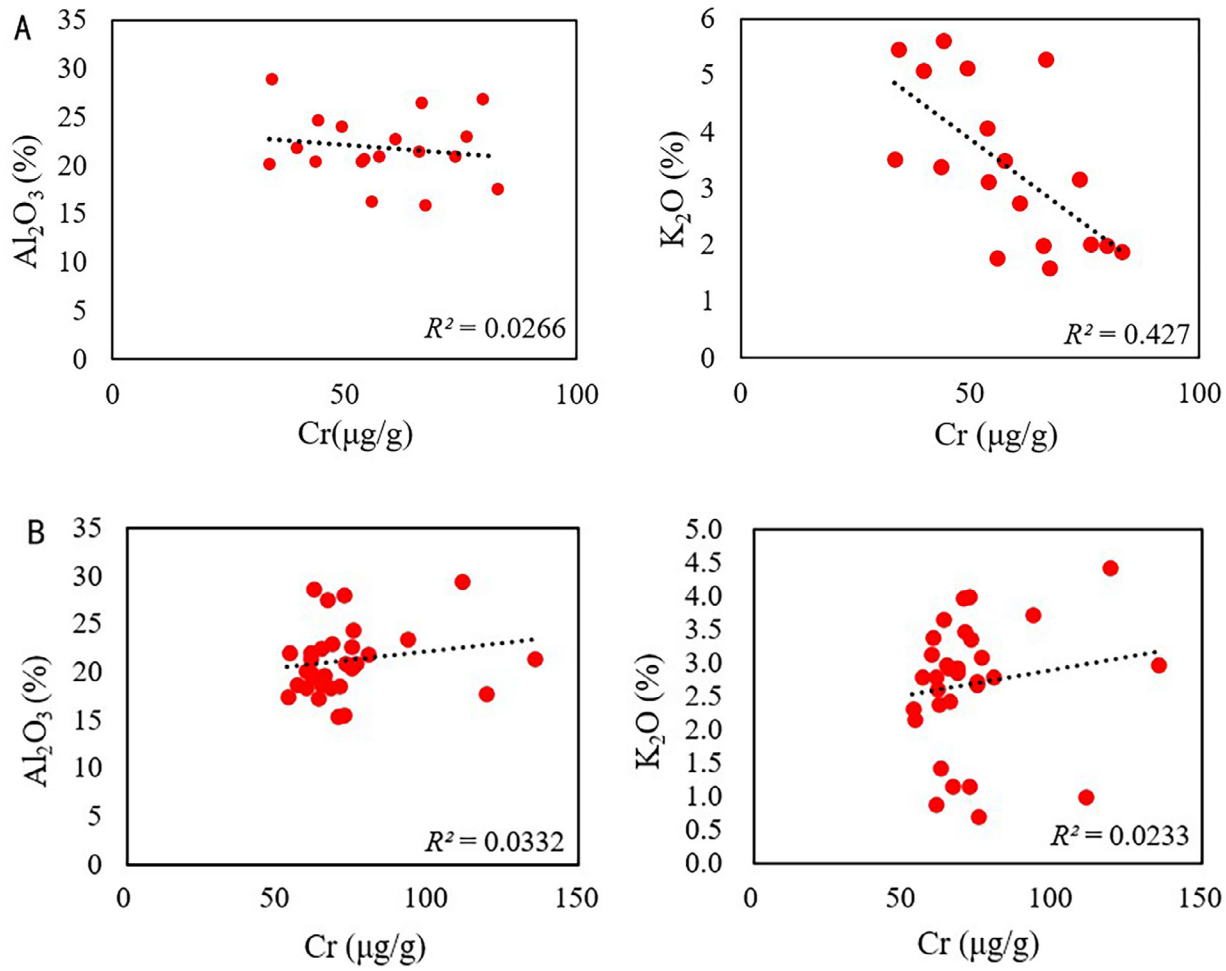



**Figure 19.** Correlation between Cr content, aluminium oxide ($Al_2O_3$), and potassium oxide ($K_2O$). (A) Taiyuan Formation coal-bearing shale. (B) Upper Paleozoic mudstone.

$$\delta Ce = \frac{Ce}{Ce^*} = \frac{Ce_N}{\sqrt{La_N \times \mathrm{Pr}_N}} \quad (2)$$

$$\delta La = \frac{La}{La^*} = \frac{La_N}{\sqrt{Ce_N \times \mathrm{Pr}_N}} \quad (3)$$

Among them, $Eu_N$, $Sm_N$, $Gd_N$, $Ce_N$, $La_N$, and $\mathrm{Pr}_N$ are the standardized values of chondrites of Eu, samarium, gadolinium, Ce, lanthanum (La), and praseodymium, respectively. Equations 1–3 calculate the abnormal Eu, Ce, La, and Y/Ho ratio in the Shiqianfeng Formation–Benxi Formation of the LX-5 well. The results are shown in Figure 21.

All samples exhibited negative Eu anomaly characteristics, which were notably different from the strong positive Eu anomaly characteristics of high-temperature hydrothermal fluids on the seabed. Therefore, the source of REE was not hydrothermal. Moreover, the light REE enrichment, Ce positive anomaly, lower Y/Ho ratio, and La positive anomaly in the samples were markedly different from the distribution characteristics of REE in seawater. Therefore, the source of REE was not from seawater or marine organisms.

### *4.6.3. Controls and enrichment mechanisms of chromium in coal-bearing strata*

The enrichment and distribution of chromium (Cr) in the study area's coal-bearing strata are not attributable to a single geological event but a result from the complex interplay of syndepositional environmental conditions and subsequent tectonic-hydrothermal overprinting. The following points outline the distinct mechanisms driving Cr accumulation, highlighting the transition from primary sedimentary controls to secondary hydrothermal

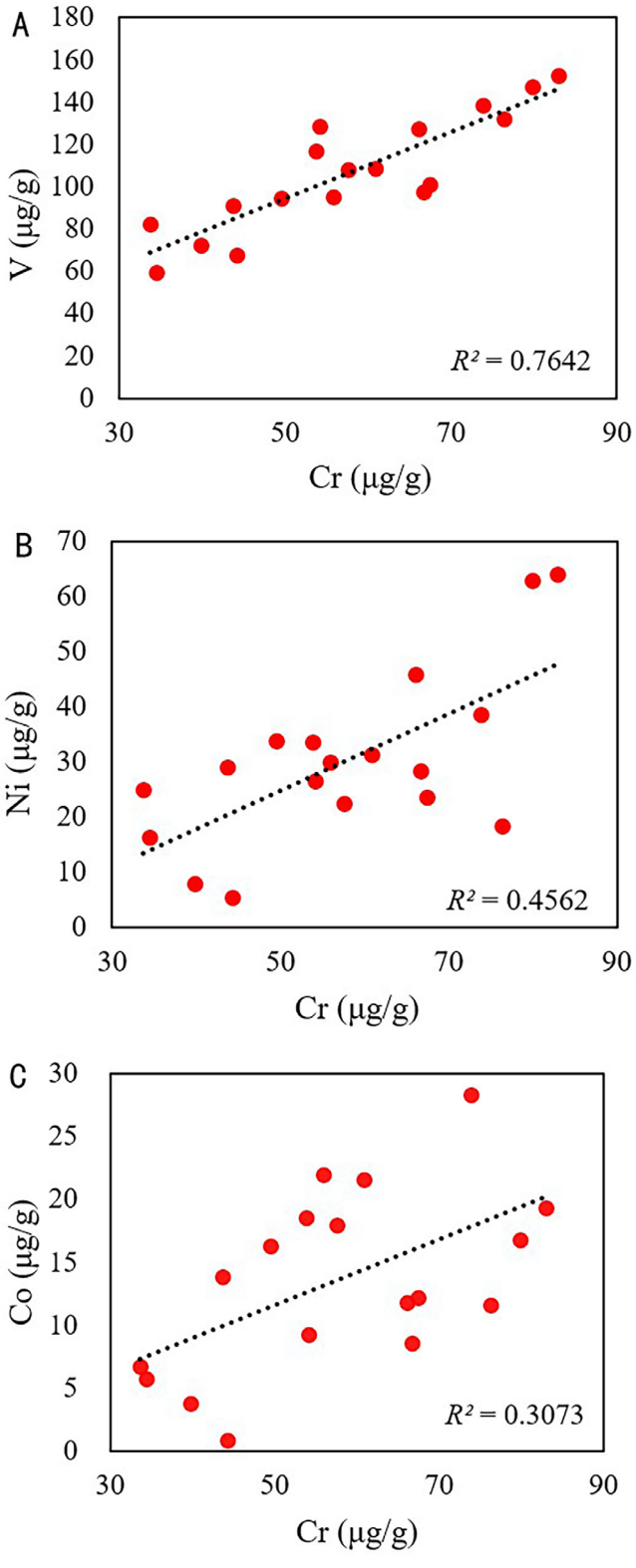



**Figure 20.** Correlation between chromium (Cr) and various elements in the lower coal-bearing shale of the Taiyuan Formation. (A) Cr and vanadium (V). (B) Cr and nickel (Ni). (C) Cr and cobalt (Co).

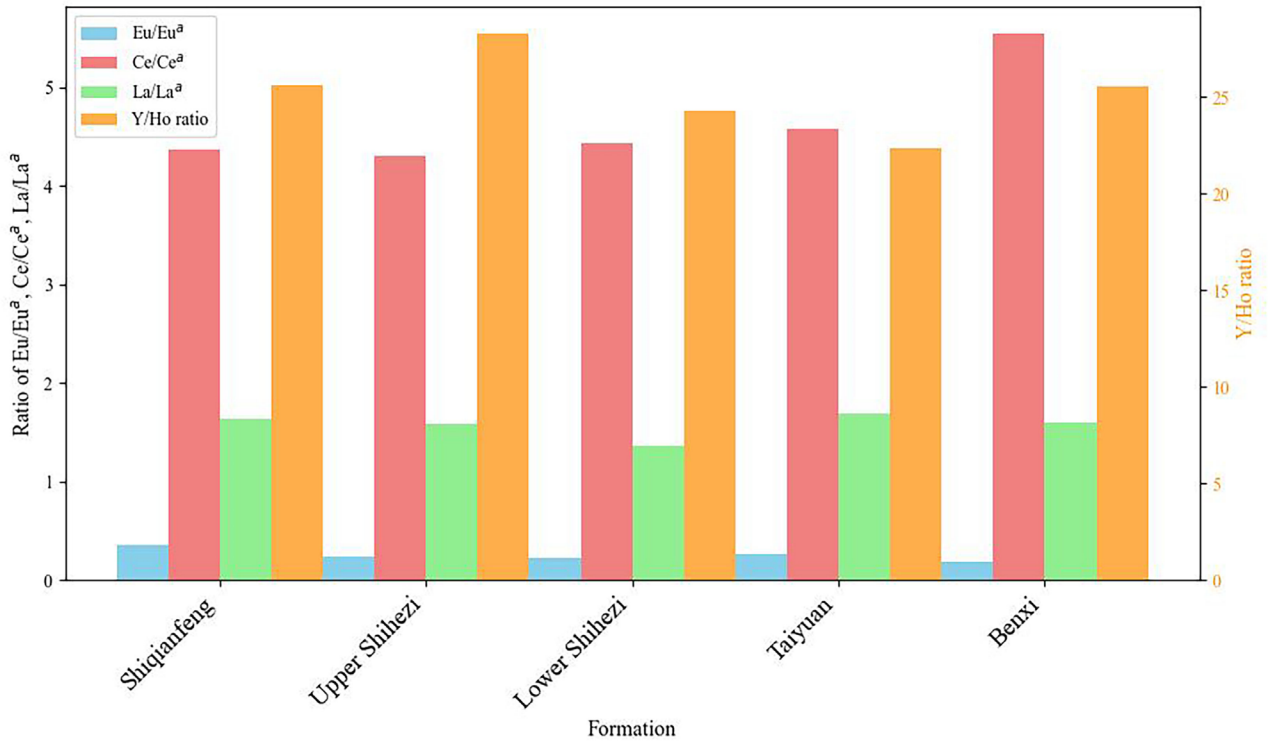



**Figure 21.** Europium (Eu), cerium (Ce), lanthanum (La) anomalies, and yttrium (Y)/holmium (Ho) ratio
Note: Eu$^a$: Theoretical Eu concentration calculated based on adjacent elements Sm and Gd; Ce$^a$: Theoretical Eu concentration calculated based on adjacent elements La and Pr; La$^a$: Theoretical Eu concentration calculated based on adjacent elements Ce and Pr.

modification and the challenges in quantifying their relative contributions.

(i) Cr element differences caused by volcanic hydrothermal activity:

The Cr element in the Paleozoic sandstone of this area showed significant stratigraphic differentiation across strata (Figure 22). The Lower Shihezi Formation exhibited extremely strong anomalies; the Taiyuan Formation had a high value and was mainly moderately enriched; and the Shanxi Formation samples generally showed regional, relatively uniform moderate enrichment. The Shiqianfeng Formation had significant fluctuations, and local high values may have multi-stage superposition or transformation characteristics. The stone box group was close to the background value (59.5–62.0 μg/g). Similarly, in the Upper Paleozoic sandstone samples, the Lower Shihezi Formation L5-18 was an outlier, with a Cr content (307.33 μg/g) close to 10 times the background value. Overall, high-value outliers showed significant locality and heterogeneity.

The differences in Figure 22 were mainly controlled by Yanshanian tectonic volcanic hydrothermal activity induced by the Zijinshan rock-mass uplift. The Taiyuan Formation, a reducing coal-bearing formation rich in organic matter and pyrite, experienced rapid Cr enrichment at fixed points. After hydrothermal fluids rose along the deep fault, precipitation and adsorption were triggered by a sudden drop in temperature and pressure and a strong reducing environment at the intersection of the fault, coal seam, and mudstone, leading to abnormal enrichment. The thick sandstone of the Lower Shihezi Formation had high pore permeability, making it an effective conduit for lateral hydrothermal fluid migration. A large fluid flux and water-rock reaction promoted widespread Cr precipitation in sandstone pores, cement, or sandstone-shale contact zones, forming a strong enrichment zone like L5-18. The Shanxi Formation may be on the periphery of the main channel, showing secondary diffusion and adsorption superposition of hydrothermal fluids during transport, resulting in a relatively stable background and moderate Cr enrichment. In contrast, the Shiqianfeng Formation lay in the uppermost layer and was less affected by deep hydrothermal fluids. Its Cr content reflected the initial background of terrestrial debris input and later surface fluid alteration, and local high values may be related to occasional fault communication.

In summary, the interlayer differences of Cr in the study area are dominated by volcanic hydrothermal events. High-value anomalies mainly develop near fault systems and are coupled with high-porosity, permeable sandstones or strongly reducing organic-rich layers. Moderate enrichment reflects the regional effect of hydrothermal diffusion superimposed on the original sedimentary background. This implies that when predicting the associated enrichment of critical metals in coal measures, priority should be given to favorable sections controlled by deep faults around the uplifted rock mass, sandstone channels of the Lower Shihezi Formation, and reducing coal-bearing traps of the Taiyuan Formation.

(ii) Cr element response and sedimentary environment: Figure 23 shows a planar distribution map of Cr element content in the lower coal-bearing shale of the Taiyuan Formation in the study area, with the red box area representing the 3D seismic study area. Based on the RMS amplitude attributes and sedimentary microfacies interpretations of the Shanxi and Taiyuan formations (Figure 7), the corresponding relationships among seismic attributes, sedimentary microfacies, and element-enrichment response were established.

The Cr content was relatively high in the northern study area, and the overall Cr level in the 3D study area was moderate to high. Based on sedimentary facies distribution, high-value Cr areas often correspond to low-energy muddy sedimentary zones. As a typical iron-clay element, Cr is easily enriched in fine-grained sediments through multiple complexation and adsorption processes. Under relatively reduced sedimentary-interface conditions, its migration loss is reduced, which is beneficial for preservation. Therefore, static or weak hydrodynamic environments, such as distributary bays, lagoons, and tidal flats, are conducive to the deposition and preservation of fine-grained materials, organic matter, and Cr enrichment in muddy sediments. In contrast, high-energy sandy microfacies are usually not conducive to Cr enrichment due to increased particle size and dilution.

In summary, Cr enrichment in the study area is jointly controlled by low-energy, fine-grained sedimentation and relative reduction preservation conditions, and its spatial distribution exhibits a strong coupling with muddy microfacies.

(iii) Relative contributions of syndepositional and

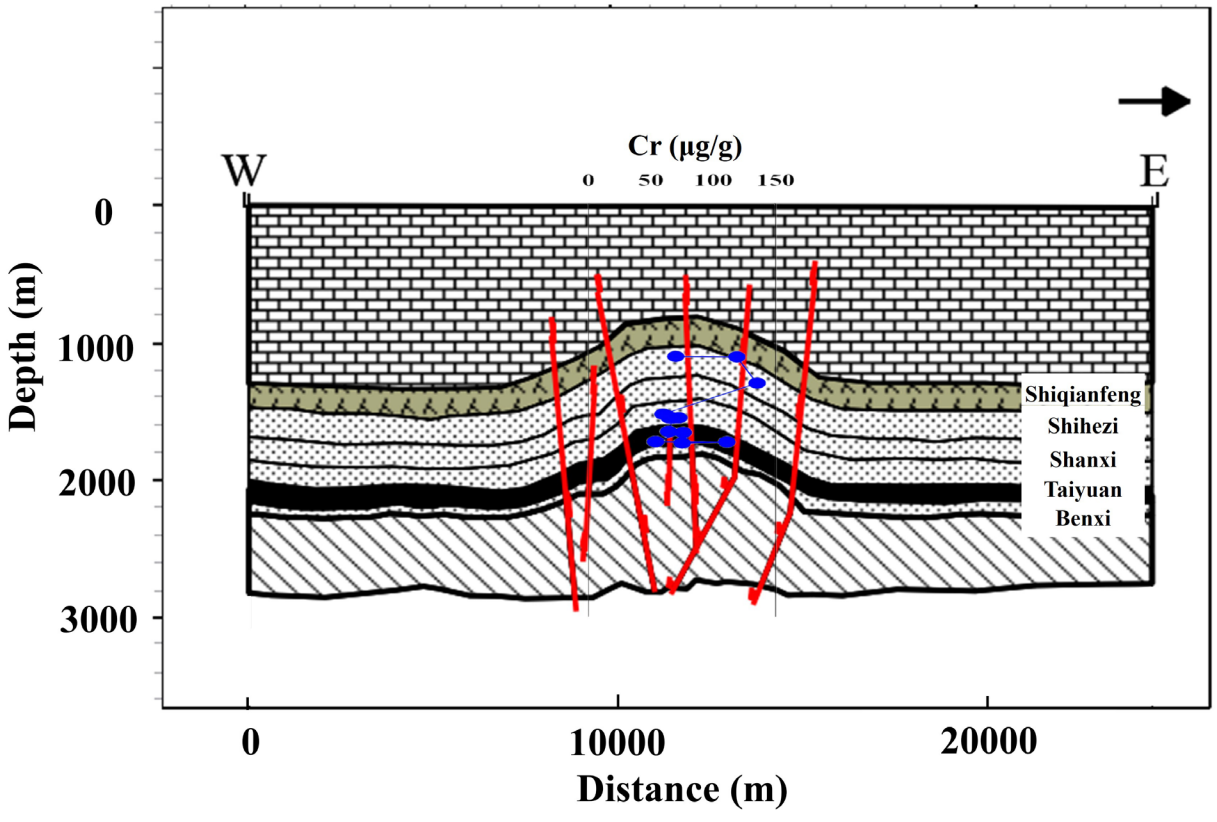



**Figure 22.** The variation of chromium (Cr) element content in Upper Paleozoic mudstone.

hydrothermal enrichment:

While this study delineates a multi-stage enrichment model for Cr, a quantitative assessment of the relative contributions from syndepositional versus post-depositional hydrothermal processes remains challenging. The anomalous Cr enrichment observed in the Lower Shihezi Formation (sample L5-18) and the localized high-value zones in the Taiyuan Formation are interpreted as resulting from synergistic effects between initial sedimentary concentration and subsequent hydrothermal remobilization. However, the overprinting of primary geochemical signatures by hydrothermal fluids, the potential mixing of multiple metal sources, and the absence of unique tracers exclusively tied to a single process currently preclude a precise quantitative partitioning. Future studies employing coupled isotopic systems (e.g., Cr, Sr, and neodymium), high-resolution micro-analytical techniques on mineral phases, and reactive transport modeling could help to disentangle these contributions. Nevertheless, the qualitative framework established here, which emphasizes the sequential roles of sedimentary reduction, fault-controlled fluid channeling, and groundwater-mediated redistribution, provides a robust foundation for identifying exploration targets where such multi-stage enrichment is most likely to occur.

#### *4.6.4. Multi-factor comprehensive mineralization*

Based on the systematic analysis, a "source transport accumulation" dynamic mineralization model was developed for key metals such as Cr in coal-bearing strata:

(i) Source stage

The key metals (such as Cr, Ga, and Zr) mainly originate from the weathering and erosion products of aluminum potassium granite and basic/ultrabasic rocks in the northern Yinshan ancient land. These sources, in the form of stable heavy minerals, such as chromite and clay minerals, are transported by rivers into sedimentary basins.

(ii) Same sedimentary enrichment stage

In the warm and humid freshwater tidal-delta barrier-island reduction environment of the Taiyuan Formation, metals are mainly adsorbed onto clay minerals or initially deposited onto heavy mineral particles. The strong reduction conditions, high organic matter content, and presence of sulfides promote the initial fixation of metals through sulfurization, organic complexation, and adsorption onto fine-grained sediments. At this stage, the chemical properties of pore water (early diagenetic groundwater) in sediments, such as oxygen deficiency and sulfur content, directly control the stability and organic complexation efficiency of metal sulfides. The geochemical gradient generated by salinity fluctuations caused by marine invasion further affects the distribution of metals between water and sediment. At this stage, a regional background enrichment pattern, constrained by stratigraphy, formed.

(iii) Postnatal transformation and secondary enrichment stage

The Yanshanian tectonic thermal event (related to the Zijinshan pluton) is the key to triggering secondary enrichment. This stage involves two interrelated fluid systems: one is magma hydrothermal fluid originating

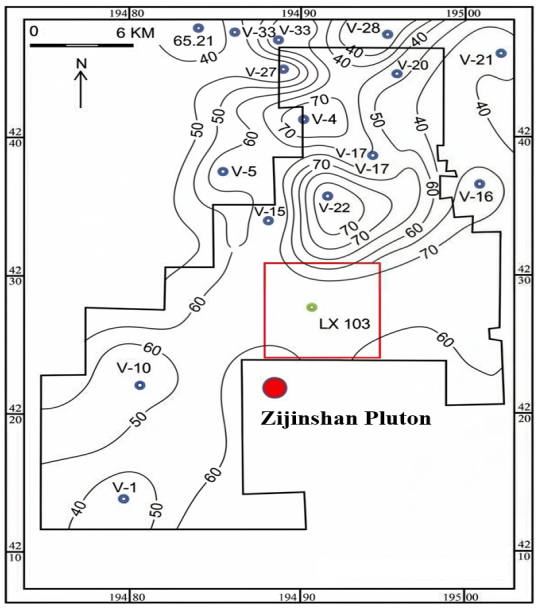



**Figure 23.** The plane distribution map of chromium content (μg/g) in coal measure shale of the Lower Taiyuan Formation

from the rock mass, upwelling along deep faults and fracture networks; The second is the groundwater in the basin that is supplied by the eastern edge of the basin and circulated under the joint control of sedimentary facies (high-permeability sandstone) and structurally-enhanced permeable zones (fault zones). The fault fissure system provides an advantageous transport pathway for both deep hydrothermal fluids and ancient groundwater. The mixing of two fluids, the formation of redox interfaces, and the long-distance transport of groundwater guided by structures collectively drive the activation and migration of fixed metals, and result in reprecipitation at favorable locations, such as fault intersections, high porosity and permeability sandstones, or reducing traps, ultimately leading to local high-grade enrichment.

## 5. Conclusion

Based on the integrated analysis of multi-source geophysical, geochemical, and geological data from the eastern margin of the Ordos Basin, this study elucidates the multi-stage enrichment mechanisms of critical metals within the coal-bearing strata.

The sedimentary environment of the Taiyuan Formation, characterized by a warm-humid, freshwater-influenced tidal delta-barrier island system, established the primary geochemical baseline. The resulting reducing and organic-rich conditions favored the initial sequestration of metals, including Zr, Cr, Ni, and Hf, through processes such as sulfidation and organo-complexation. Subsequent Yanshanian tectono-thermal events, associated with the Zijinshan pluton, were pivotal drivers of secondary enrichment. The magmatic activity generated fault and fracture networks that served as primary conduits for metal-bearing hydrothermal fluids, leading to localized, high-grade enrichment within high-porosity sand bodies and fault zones. Groundwater circulation played a critical and previously underappreciated role as a mobilizing agent. Recharged from the eastern basin margin, paleo-groundwater flow, regulated by sedimentary facies and fracture systems, facilitated the remobilization, redistribution, and reprecipitation of metals in response to hydrological and redox gradients. A focused study on Cr enrichment delineates a representative three-stage model: source weathering (terrigenous input), sedimentary reduction (initial fixation), and post-depositional tectonic-hydrothermal-groundwater activation (secondary enrichment). This model underscores the coupled influence of provenance, depositional setting, and post-depositional fluid activity.

The structure–sedimentation–groundwater coupled joint ore-controlling model established in this study deepens and enriches the previous conceptual framework regarding the control of minerals by single geological processes. It also clarifies the crucial bridging and core driving functions of ancient groundwater systems in this context. In contrast to previous studies that have predominantly focused on static reservoirs or isolated hydrothermal systems, this paper reveals the spatial framework of dominant groundwater transport pathways through high-precision wave impedance inversion and 3D attribute–structure coupling modeling. Moreover, by comprehensively leveraging geochemical indicators such as the Th/U ratio, it dynamically traces the oxidation-leaching activation ability and metal transport capacity. This model offers a more systematic and dynamic elucidation of the key metal mineralization mechanisms in coal-bearing strata and analogous basin environments. It also provides explicit predictive guidance for resource exploration in these areas.

## Acknowledgments

None.

## Funding

This study is supported by the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (No. 42321002).

## Conflict of interest

The authors declare no conflicts of interest.

## Author contributions

*Conceptualization:* Yi Yang, Jingtao Zhao
*Formal analysis:* Yi Yang, Wanli Gao
*Funding acquisition*: Jingtao Zhao
*Investigation:* Suzhen Shi
*Methodology:* Yi Yang, Jingtao Zhao
*Writing–original draft:* Yi Yang
*Writing–review & editing:* Jingtao Zhao, Suzhen Shi

## Availability of data

The data used in this study are restricted due to privacy and confidentiality agreements with the data providers. Data are available from the corresponding author upon reasonable request.

## References

1. Ning SZ, Huang SQ, Liu K, *et al*. Comparison of genesis of abnormal enrichment of metals in coal between the northern and southern margins of Ordos Basin. *J China Coal Soc.* 2022;47(5):1795-1807.

doi: 10.13225/j.cnki.jccs.MJ22.0367

2. Cao B, Fu XH, Kang JQ, Tang P, Xu H. Enrichment Factors and Metallogenic Models of Critical Metals in Late Permian Coal Measures from Yunnan, Guizhou, and Guangxi Provinces. *Minerals*. 2024;14(2):206.

doi: 10.3390/min14020206

3. Cao DY, Qin GH, Zhang Y, *et al*. Classification and combination relationship of mineral resources in coal measures. *J China Coal Soc*. 2016;41(9):2150-2155.

doi: 10.13225/j.cnki.jccs.2016.0776

4. Zhao L, Dai SF, Nechaev VP, Nechaeva EV, Graham IT, French D. Enrichment origin of critical elements (Li and rare earth elements) and a Mo-U-Se-Re assemblage in Pennsylvanian anthracite from the Jincheng Coalfield, southeastern Qinshui Basin, northern China. *Ore Geol Rev.* 2019;115:103184.

doi: 10.1016/j.oregeorev.2019.103184

5. Zhao L, Wang XB, Dai SF. Lithium resources in coal-bearing strata: Occurrence, mineralization, and resource potential. *J China Coal Soc*. 2022;47(5):1750-1760.

doi: 10.13225/j.cnki.jccs.MJ22.0418

6. Ren DY, Zhao FH, Zhang JY, Xu D. A Preliminary study on genetic type of enrichment for hazardous minor and trace elements in coal. *Earth Sci Front.* 1999;6(z1):17-22.

doi: 10.3321/j.issn:1005-2321.1999.z1.003

7. Liu JJ, Dai SF, Song HJ, *et al*. Geological factors controlling variations in the mineralogical and elemental compositions of Late Permian coals from the Zhijin-Nayong Coalfield, western Guizhou, China. *Int J Coal Geol.* 2021;247:103855.

doi: 10.1016/j.coal.2021.103855

8. Dai SF, Wang PP, Ward CR, *et al*. Elemental and mineralogical anomalies in the coal-hosted Ge ore deposit of Lincang, Yunnan, southwestern China: Key role of N2–CO2-mixed hydrothermal solutions. *Int J Coal Geol*. 2015;152:19-46.

doi: 10.1016/j.coal.2014.11.006

9. Wei Q, Dai SF. Critical metals and hazardous elements in the coal-hosted germanium ore deposits of China: Occurrence characteristics and enrichment causes. *J China Coal Soc*. 2020;45(1):296-303.

doi: 10.13225/j.cnki.jccs.YG19.1633

10. Liu JJ, Han QC, Zhao SM, Jia R. The sources of abnormally enriched critical metals in the Late Permian coals of Western Guizhou Province. *J China Coal Soc*. 2022;47(5):1782-1794.

doi: 10.13225/j.cnki.jccs.MJ22.0178

11. Cao DY, Wei YC, Qin GH, *et al*. Tectonic control on enrichment and metallogenesis of strategic metal elements in coal measures. *Coal Geol Explor.* 2023;51(1):66-85.

doi: 10.12363/issn.1001-1986.22.09.0734

12. Zhang XL, He JX, Dong SH, Huang YP, Shi S, Wang W. Enrichment Characteristics and Genetic Analysis of the Strategic Metal Element Gallium in Coal-measure Shale, Lower Taiyuan Formation, Linxing Area. *Acta Sedimentol Sin.* 2022;40(6):1676-1690.

doi: 10.14027/j.issn.1000-0550.2022.110

13. Wu SG, Yu ZH, Zhang RQ, Han WG, Zou DB. Mesozoic–Cenozoic tectonic evolution of the Zhuanghai area, Bohai-Bay Basin, east China: the application of balanced cross-sections. *J Geophys Eng.* 2005;2(2):158-168.

doi: 10.1088/1742-2132/2/2/011

14. Sun QL, Wu SG, Lü FL, Yuan SQ. Polygonal faults and their implications for hydrocarbon reservoirs in the southern Qiongdongnan Basin, South China Sea. *J Asian Earth Sci.* 2010;39(5):470-479.

doi: 10.1016/j.jseaes.2010.04.002

15. Jamaludin SNF, Latiff AHA, Ghosh DP. Structural balancing vs horizon flattening on seismic data: Example from extensional tectonic setting. In: Gaol FL, Soewito B, editors. *IOP conference series: Earth and environmental science*. 2nd International Conference on Geological, Geographical, Aerospace and Earth Sciences 2014 (AeroEarth 2014); October 11–12, 2014; Bali, Indonesia. IOP Publishing; 2015;23(1):012003.

doi: 10.1088/1755-1315/23/1/012003

16. Budach I, Moeck I, Lüschen E, Wolfgramm M. Temporal evolution of fault systems in the Upper Jurassic of the Central German Molasse Basin: case study Unterhaching. *Int J Earth Sci.* 2018;107(2):635-653.

doi: 10.1007/s00531-017-1518-1

17. Quan H, Yu Z, Qiao Z, *et al*. Sedimentary and Early Diagenetic Responses to the Huaiyuan Movement During the Early–Middle Ordovician Transition in the Ordos Basin, North China. *Geosciences*. 2025;15(6):219.

doi: 10.3390/geosciences15060219

18. Qin GH, Deng LJ, Liu K, Xu H, Ma ZK, Cao DY. Characteristic of rare earth elements in coal in western margin of Ordos basin. *Coal Geol Explor.* 2016;44(6):8-14.

doi: 10.3969/j.issn.1001-1986.2016.06.002

19. Zhou XQ, Qin Y, Lu L. Advances on geological-geochemical research of coal-type uranium in China. *Coal Geol Explor*. 2019;47(4):45-53.

doi: 10.3969/j.issn.1001-1986.2019.04.008

20. Wang JQ, Deng JX, Xia H, Yan LL. Rock physics characteristics and their control factors of carbonate in different sedimentary microfacies of the Yingshan Formation, Gucheng Area, Tarim Basin. *J Geophys Eng.* 2023;20(6):1308-1321.

doi: 10.1093/jge/gxad087

21. Li XT, Lin CS. Delta front depositional system study based

on seismic attributes and well logging facies. *Appl Mech Mater.* 2012;170-173:1285-1289.

doi: 10.4028/www.scientific.net/AMM.170-173.1285

22. Wang Z, Xian BZ, Ma LC, Chao CZ, Pu Q, Liu JP. Mapping sediment-dispersal characteristics of Neogene deltas using seismic geomorphology, Liaodongdong area, Bohai Bay Basin, China. *Energy Sources Part A Recovery Util Environ Eff.* 2018;40(19):2265-2276.

doi: 10.1080/15567036.2018.1486477

23. Cao L, Chang S, Yao Y. Applying seismic sedimentology in predicting sedimentary microfacies and coalbed methane gas content. *J Nat Gas Sci Eng.* 2019;69:102944.

doi: 10.1016/j.jngse.2019.102944

24. Liu SM, Zhou L, Guan X, *et al*. Research on distribution characteristics of sedimentary microfacies of a system tract under a high resolution sequence framework: A case study of Qixia Formation in Gaomo block, central Sichuan. *Front Earth Sci.* 2023;11:1101242.

doi: 10.3389/feart.2023.1101242

25. Hower JC, Ruppert LF, Eble CF. Lanthanide, yttrium, and zirconium anomalies in the fire Clay coal bed, Eastern Kentucky. *Int J Coal Geol.* 1999;39(1-3):141–153.

doi: 10.1016/S0166-5162(98)00043-3

26. Mukhopadhyay PK, Goodarzi F, Crandlemire A.L, Gillis KS, MacNeil DJ, Smith WD. Comparison of coal composition and elemental distribution in selected seams of the Sydney and Stellarton Basins, Nova Scotia, Eastern Canada. *Int J Coal Geol*. 1998;37(1–2):113–141.

doi: 10.1016/S0166-5162(98)00020-2

27. Seredin VV, Finkelman RB. Metalliferous coals: a review of the main genetic and geochemical types. *Int J Coal Geol*. 2008;76(4):253–289.

doi: 10.1016/j.coal.2008.07.016

28. Dai SF, Ren DY, Li SS, Chou C. A Discovery of Extremely-enriched Boehmite from Coal in the Junger Coalfield, the Northeastern Ordos Basin. *Acta Geol. Sin*. 2006;80(02):294-300+315-316.

doi: 10.3321/j.issn:0001-5717.2006.02.015

29. Zou JH. Critical Elements Enrichment Mechanism of the Lopingian Coal-bearing Strata in the Eastern Sichuan Basin. PhD Thesis, China University of Mining and Technology (Beijing), Beijing; 2018.

30. Li Y, Wu P, Gao JX, *et al*. Multilayer coal-derived gas enrichment mechanism and whole gas-bearing system model: A case study on the Linxing block along the eastern margin of the Ordos Basin. *Nat Gas Ind*. 2022;42(6):52-64.

doi: 10.3787/j.issn.1000-0976.2022.06.005

31. Wu P, Gao LJ, Li Y, *et al*. An evaluation method for shale gas potential of marine-continent transitional facies with frequent interbedded lithology: A case study on the Lower Permian Shanxi Formation in Linxing block of the Ordos Basin. *Nat Gas Ind*. 2022;42(2):28-39.

doi: 10.3787/j.issn.1000-0976.2022.02.004

32. Du J, Zhu GH, Li Y, Wu P, Gao JX, Zhu YH. Exploration and development challenges and technological countermeasures for tight sandstone gas reservoirs in Ordos basin margin: A case study of Linxing-Shenfu gas field. *Nat Gas Ind*. 2022;42(1):114-124.

doi: 10.3787/j.issn.1000-0976.2022.01.011

33. Yue H, Guo P, Zhang T, *et al*. Integrated geophysical exploration of favorable potassium salt mineralization on the northeastern margin of Qaidam Basin, China. *J Geophys Eng.* 2025;22(3):889-900.

doi: 10.1093/jge/gxaf048

34. Li Y, Xu WK, Gao JX, *et al*. Mechanism of coal measure gas accumulation under integrated control of "source reservoir-transport system": A case study from east margin of Ordos Basin. *J China Coal Soc.* 2021;46(8):2440-2453.

doi: 10.13225/j.cnki.jccs.CB21.0765

35. Guo Q, Wang XY, Yin S, Wu ZH. Tectonic evolution and gas control analysis of the Upper Palaeozoic tight sandstone reservoirs in the Linxing area of the eastern Ordos Basin. *Geol. J.* 2021;56(3):1625-1637.

doi: 10.1002/gj.4007

36. Dai SF, Seredin VV, Ward CR, *et al*. Enrichment of U-Se-Mo-Re-v in coals preserved within marine carbonate successions: geochemical and mineralogical data from the Late Permian Guiding Coalfield, Guizhou.China. *Miner Depos.* 2015;50(2):159-186.

doi: 10.1007/s00126-014-0528-1

37. Fan YH, Qu HJ, Wang H, Yang XC. The application of trace elements analysis to identifying sedimentary media environment：A case study of Late Triassic strata in the middle part of western Ordos Basin. *Geol. China.* 2012;39(2):382-389.

doi: 10.3969/j.issn.1000-3657.2012.02.010

38. Jones B, Manning DAC. Comparison of geochemical indices used for the interpretation of palaeoredox conditions in ancient mudstones. *Chem Geol.* 1994;111(1-4):111-129.

doi: 10.1016/0009-2541(94)90085-x

39. Rimmer, SM. Geochemical paleoredox indicators in Devonian-Mississippian black shales, Central Appalachian Basin (USA). *Chem Geol*. 2004;206(3-4):373–391.

doi: 10.1016/j.chemgeo.2003.12.029

40. Shi J, Huang WH, Lv CH, Cui X. Geochemical characteristics and geological significance of the Upper Paleozoic mudstones from Linxing area in Ordos Basin. *Acta Pet Sin.*

2018;39(8):876-889.

doi: 10.7623/syxb201808004

41. Wu HC, He JX, Zhang XL, Ren ZQ, Zhou TT, Wang AK. Pore structure and its fractal features of the shale in Taiyuan Formation of Linxing area, Ordos Basin. *Xinjiang Pet Geol.* 2018;39(5):549-554.

doi: 10.7657/XJPG20180508

42. Tribovillard N, Algeo TJ, Lyons T, Riboulleau A. Trace metals as paleoredox and paleoproductivity proxies: an update. *Chem Geol*. 2006;232(1-2):12-32.

doi: 10.1016/j.chemgeo.2006.02.012

43. Dai SF, Ren DY, Chou CL, Finkelman RB, Seredin VV, Zhou YP. Geochemistry of trace elements in Chinese coals: A review of abundances, genetic types, impacts on human health, and industrial utilization. *Int J Coal Geol.* 2012;94:3-21.

doi: 10.1016/j.coal.2011.02.003

44. Finkelman RB, Bostick NH, Dulong FT, Senftle FE, Thorpe AN. Influence of an igneous intrusion on the inorganic geochemistry of a bituminous coal from Pitkin County, Colorado. *Int J Coal Geol*. 1998;36(3-4):223–241.

doi: 10.1016/S0166-5162(98)00005-6

45. Stewart AK, Massey M, Padgett PL, Rimmer SM, Hower JC. Influence of a basic intrusion on the vitrinite reflectance and chemistry of the Springfield (No. 5) coal, Harrisburg, Illinois. *Int J Coal Geol*. 2005;63(1-2):58–67.

doi: 10.1016/j.coal.2005.02.005

46. Sun R, Liu G, Zheng L, Chou CL. Geochemistry of trace elements in coals from the Zhuji Mine, Huainan Coalfield, Anhui, China. *Int J Coal Geol*. 2010;81(2):81–96.

doi: 10.1016/j.coal.2009.12.001

47. Russell BH. *Introduction to Seismic Inversion Methods*. Tulsa, OK, USA: Society of Exploration Geophysicists (SEG). 1988.

48. Yilmaz, Ö. *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Tulsa, OK, USA: Society of Exploration Geophysicists (SEG). 2001.

49. Azizzadeh Mehmost Olya B, Mohebian R, Moradzadeh A. A new approach for seismic inversion with GAN algorithm. *J Seism Explor.* 2024,33(3):01–36.

50. Zhang Y, Singh S, Thanoon D, Thanoon D, Devarakota P, Jin L, Tsvankin I. Physics-guided unsupervised deep-learning seismic inversion with uncertainty quantification. *J Seism Explor*. 2023;32(3):257–270.

51. Wang P, Xu HQ, Peng Z, Wang ZF, Yang MQ. Application of data augmentation based on generative adversarial network in impedance inversion. *J Seism Explor*. 2023;32(2):155–168.

52. Cai YD, Gao GS, Liu DM, Qiu F. Geological conditions for coal measure gas enrichment and accumulation models in Linxingzhong Block along the eastern margin of the Ordos Basin. *Nat Gas Ind.* 2022;42(11):25-36.

doi: 10.3787/j.issn.1000-0976.2022.11.003

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Automatic differentiation-based weighted instantaneous phase inversion

**Mengzi Wang, Yong Hu*, Zhihan Zhang, Chunyuan Shi, and Xiaoyuan Liao**

Department of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu, China

(This article belongs to the Special Issue: *Full Waveform Inversion Methods and Applications for Seismic Data in Complex Media*)

## Abstract

Traditional full waveform inversion objective functions typically rely on discrepancies between waveforms, making full waveform inversion highly dependent on the initial model and the low-frequency components in the seismic data. The instantaneous phase of the seismic signal reflects the kinematic information of seismic waves and has a more linear relationship with the subsurface velocity structure, aiding in the retrieval of low-wavenumber components of velocity models. Conversely, waveform information is instrumental in achieving high-resolution inversion results while maintaining algorithmic stability. In addition, automatic differentiation provides an efficient and accurate mechanism for computing gradients by systematically applying the chain rule across the computational graph, yielding reliable derivative information for optimization. To balance the contributions of waveform and phase information in velocity inversion, within the framework of automatic differentiation, we integrate instantaneous phase and waveform information to propose an automatic differentiation-based weighted instantaneous phase inversion method. Numerical tests using the Marmousi model and the Overthrust model demonstrate that the proposed method achieves more accurate velocity inversion results.

***Keywords:*** Full waveform inversion; Instantaneous phase; Weighted objective function; Automatic differentiation; Cycle skipping

***Corresponding author:**
Yong Hu
(jluhuyong@sina.com)

**Citation:** Wang M, Hu Y, Zhang Z, Shi C, Liao X. Automatic differentiation-based weighted instantaneous phase inversion. *J Seismic Explor.* 2026;35(2):025520132. doi: 10.36922/JSE025520132

**Received:** December 25, 2025

**Revised:** January 28, 2026

**Accepted:** January 29, 2026

**Published online:** March 27, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

In recent years, driven by the continuous improvement of computing capabilities, full waveform inversion (FWI) has gradually become a research hotspot. FWI leverages complete recorded seismic waveforms, iteratively minimizing the objective function that measures the misfit between synthetic and observed data to invert the velocity parameters of subsurface structures. Since its introduction in 1983, FWI has been recognized as a vital tool for high-fidelity subsurface physical parameter estimation.[1,2] Traditional FWI objective functions are typically formulated using the L2 norm, and gradients are calculated using the adjoint-state method to guide model updates in a descent direction.[3] However, in the absence of low-frequency components, directly matching synthetic and observed waveforms can lead to cycle-skipping, causing the

objective function to converge to local minima.[4-10]

Compared with waveform information, phase information in seismic data mainly reflects the kinematic attributes of seismic waves and generally has a more linear relationship with the subsurface velocity structure.[11] However, the application of phase inversion methods is constrained by the phase-wrapping issue; to effectively utilize phase information, new strategies must be employed to address this challenge. Shah *et al.*[12] and Choi and Alkhalifah[13] obtained instantaneous travel-time information by computing the derivative of phase with respect to angular frequency in the frequency domain, applied this information to inversion, and thereby effectively mitigated the impact of phase wrapping. These methods leverage the frequency-domain complex wavefield's natural separation of amplitude and phase. Nevertheless, the instantaneous phase can also be extracted and used directly in the time domain. For example, Luo *et al.*[14] proposed a time-domain acoustic FWI using instantaneous phase and demonstrated its favorable properties. Hu *et al.*[15] proposed the phase cross-correlation algorithm, which further mitigates the impact of amplitude information on inversion results. To obtain a robust initial velocity model for FWI, phase-based seismic inversion can effectively mitigate the cycle-skipping issue; examples include adaptive FWI and phase correction.[16-18]

Computing the gradient of the objective function has traditionally depended on techniques such as manual differentiation, symbolic differentiation, and numerical differentiation. These approaches, however, are plagued by inherent limitations such as cumbersome derivations, expression swelling, and truncation errors. To address these drawbacks, this study employs automatic differentiation (AD), which is specifically designed for gradient computation. AD enables automatic gradient extraction for complex differentiable functions,[19] serving as the cornerstone of modern neural network training[20] and is more commonly referred to as backpropagation.[21,22] While the concept of AD is not new, its foundational ideas were first proposed in the 1950s by Nolan,[23] who pioneered the quest to automatically differentiate computer programs. A significant milestone was reached in 1980 when Speelpenning[24] developed the first fully automated implementation of reverse-mode AD. Building on this foundation, Kasim and Vinko[25] made further advancements by developing the ξ-torch library within the PyTorch framework. This library provides differentiable scientific simulation modules, including root finders and initial-value problem solvers, thereby facilitating the seamless integration of deep learning and scientific simulation in fields such as FWI.

In the field of seismology, early research on AD was constrained by the limited maturity of available methods and thus primarily focused on relatively simple computational problems such as ray tracing and receiver function analysis. For example, Sambridge *et al.*[26] utilized AD tools such as the Transformation of Algorithms in Fortran and the Tangent Linear and Adjoint Model Compiler to implement forward gradient calculations for these tasks and demonstrated the efficiency of reverse-mode AD in computing gradients for Gibbs free energy minimization problems. With the maturation of AD tools, more studies began to address increasingly complex waveform inversion tasks, promoting the widespread application of this technology in seismology. Subsequently, a series of studies applied AD to different aspects of FWI: Tan *et al.*[27] attempted to use AD to compute the Hessian matrix in FWI; Liao[28] built an inversion workflow for the 2D acoustic wave equation based on AD tools; Richardson[29] reformulated the traditional adjoint-state method for gradient computation in seismic FWI as a recurrent neural network, combining it with AD to improve computational efficiency and convergence speed; Cao and Liao[30] implemented a borehole FWI workflow using the AD tool TAPENADE and verified its feasibility with synthetic data. Zhu *et al.*[31] further derived and validated the theoretical equivalence between gradients obtained via AD and those computed using the adjoint-state method, and developed a seismic inversion framework based on the Julia programming language. In addition, existing studies have begun to integrate AD methods with deep learning neural networks to enhance the performance of FWI.

In this study, we first introduce the concept and advantages of AD. Next, we derive the procedure for extracting exponential phase information. Then, we explain how to introduce weighting factors to construct the weighted instantaneous phase objective function. Finally, the Overthrust and Marmousi models are used to validate the effectiveness of the weighted instantaneous phase inversion method.

## 2. Methods

### 2.1. Automatic differentiation

In FWI, differentiation is used to compute the gradient of the objective function with respect to the model parameters, enabling iterative optimization. Traditional methods of derivative computation include manual, symbolic, and numerical differentiation. Manual differentiation is labor-intensive and prone to errors; symbolic differentiation can result in expression swelling; and numerical differentiation—such as finite differences—is straightforward but introduces truncation and round-off

errors, which can accumulate and potentially compromise the stability and accuracy of the inversion process.

To overcome these limitations, we introduce AD, a technique widely used in machine learning for efficient and precise derivative computation. AD evaluates derivatives by decomposing a complex function into a series of basic operations and systematically applying the chain rule to accumulate derivative information during numerical computation. Unlike the traditional adjoint-state method, which requires analytical derivation of adjoint sources and adjoint wavefields, AD automatically records the computational graph and tracks derivative propagation throughout the forward computation. This approach eliminates the need for manual derivation of adjoint terms and reduces the risk of implementation errors.[32,33]

As shown in Figure 1, the adjoint source is a key component of adjoint-state–based FWI. When the objective function becomes complex, obtaining and implementing the adjoint source and adjoint wavefield can be difficult and error-prone. In particular, after the seismic signal is transformed into the time–frequency domain or the frequency–wavenumber domain, the objective function becomes more complex, and it is increasingly difficult to obtain the adjoint source. By contrast, AD can automatically record the computational graph and track the flow of derivatives during the forward computation, eliminating the need for manual derivation of adjoint sources and adjoint wavefields. It provides an efficient tool for the implementation of FWI with complex objective functions and multi-parameter inversion.

Automatic differentiation offers two implementation strategies: forward mode and reverse mode. Within the forward mode, derivative information flows from the input nodes to the output nodes across the computation graph. This approach is particularly efficient when dealing with a limited number of input parameters. Conversely, reverse mode propagates gradients from the output and is more suitable for situations where there are many input parameters. Given that FWI usually involves a high-dimensional model parameter space with a scalar or low-dimensional objective function, reverse-mode AD is generally preferred for efficiently and accurately computing gradients in multi-parameter FWI problems.

### 2.2. Instantaneous phase information extraction

In digital signal processing, there is a certain relationship between the real and imaginary parts of a complex signal, which is described by the Hilbert transform. The Hilbert transform is a mathematical tool that can extract the instantaneous amplitude and phase of a signal and is often used to analyze the signal's instantaneous amplitude and phase characteristics. By applying the Hilbert transform to the signal, the transformed result is treated as the imaginary part, and the original signal is taken as the real part. The analytic signal $u_a(t)$ can be constructed using **Equation 1**:

$$u_a(t) = u(t) + i \cdot u_H(t) = E(t) e^{i\varphi(t)} \tag{1}$$

where $E(t)$ is the envelope of the signal, $u(t)$ represents the received seismic signal, $u_H(t)$ represents the Hilbert transform of the signal, and $\varphi(t)$ represents the instantaneous phase information of the signal. $\varphi(t)$ can be calculated using **Equation 2**:

$$\varphi(t) = arctan \frac{u_H(t)}{u(t)} \tag{2}$$

The instantaneous phase of a signal not only reflects the travel-time information of seismic waves but also directly reveals the medium's velocity and kinematic characteristics. By incorporating phase information into the objective function, the nonlinear behavior of the objective function in FWI can be significantly improved, which helps mitigate the risk of becoming trapped in local minima. This feature makes instantaneous phase a powerful tool in FWI applications. Figure 2A shows the observed single-shot record, whereas Figure 2B and 2C show single-shot records of envelope data and instantaneous phase data, respectively.

To avoid the impact of phase wrapping on the inversion, the exponential phase information of the signal is utilized:

$$e^{i\varphi(t)} = \frac{u_a(t)}{E(t)} = \frac{u(t) + i \cdot u_H(t)}{\sqrt{u^2(t) + u_H^2(t)}} \tag{3}$$

The exponential-phase representation effectively removes instantaneous amplitude information, such that the resulting complex exponential phase primarily carries the essential phase-related characteristics of the wavefield. Moreover, computing the exponential phase avoids numerically sensitive operations such as phase unwrapping, which can produce phase jumps; this helps mitigate phase wrapping issues. In this work, the exponential-phase measurements are employed as the objective function for a weighted instantaneous phase inversion within the AD framework. This combination leverages the kinematic robustness of seismic data together with AD's exact gradient computation to improve inversion

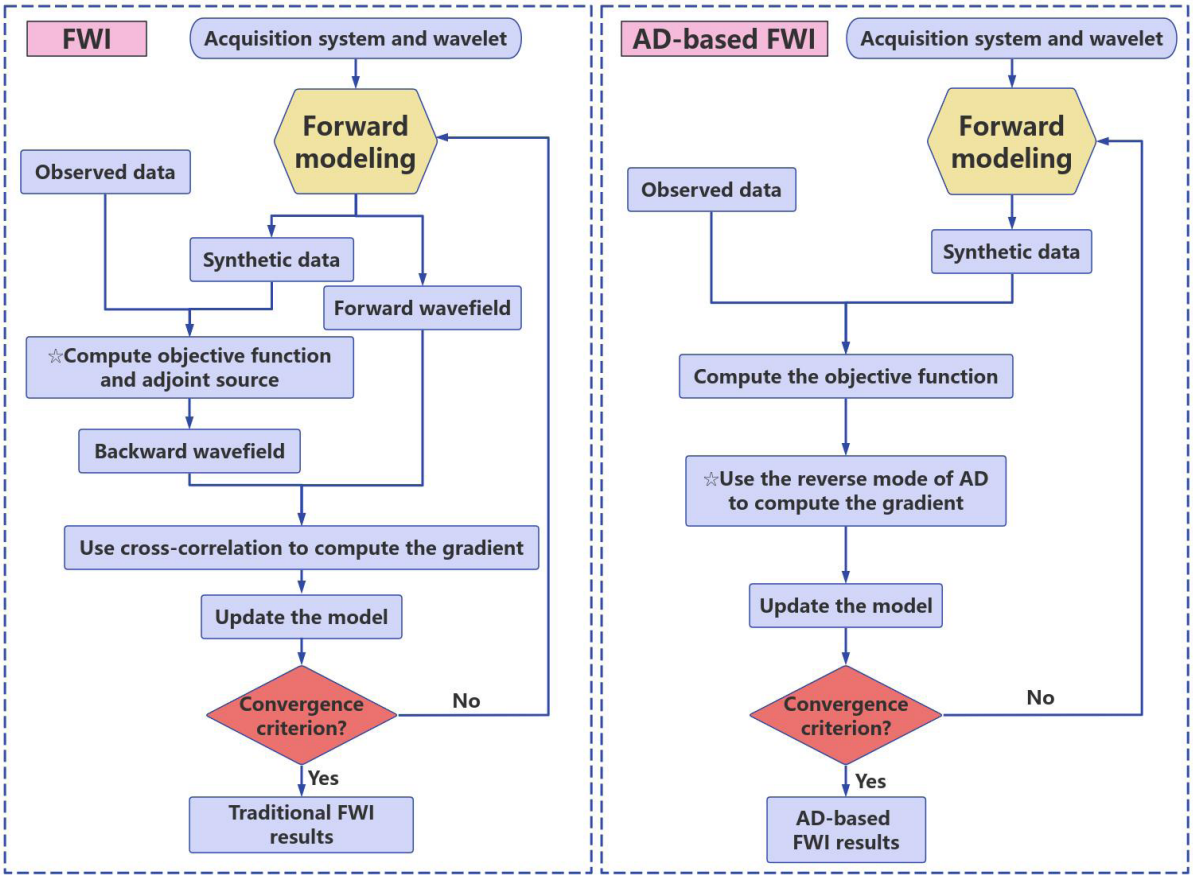



**Figure 1.** Comparison of workflows between traditional full waveform inversion (FWI) and automatic differentiation (AD)-based FWI

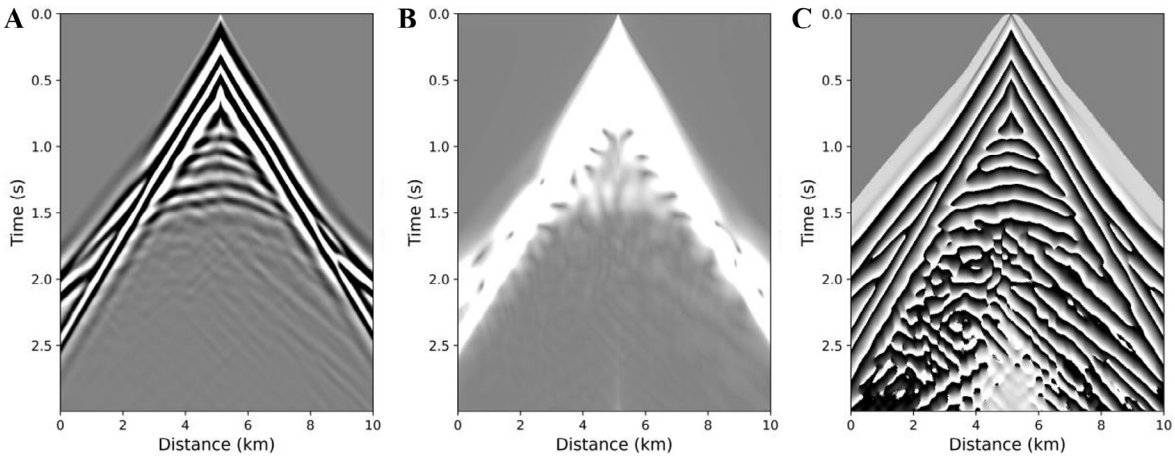



**Figure 2.** Single-shot seismic records: (A) Observed data, (B) envelope data, and (C) instantaneous phase data

stability and accuracy.

### 2.3. Construction of weighted instantaneous phase objective function

Phase information typically has a more linear relationship with the subsurface velocity structure than waveform amplitudes do. When the initial model is far from the true model and the amplitude is large, phase information more robustly captures timing/phase alignment and provides a better starting model. Conversely, when the inverted model aligns closely with the true model and fine-scale details are of interest, employing an L2 norm objective function is beneficial. This approach facilitates high-resolution inversion by fully leveraging waveform information, thereby enabling a clear depiction of complex model features and maintaining the stability of the algorithm.

Accordingly, we introduce a weighting factor, $w(i)$, to achieve a balanced integration of both instantaneous phase and waveform information. By assigning distinct weights to the objective functions based on waveform and phase information during the inversion process, the complementary strengths of both terms are utilized. The formulation for the weighted instantaneous phase objective function is expressed as follows:

$$J_{w-phase}(m) = w(i)J_w(m) + (1-w(i))J_{phase}(m) \tag{4}$$

In **Equation 4**, $J_w$ is the waveform-based L2 norm and can be expressed as:

$$J_w(m) = \frac{1}{2}\sum_s\sum_r\int_0^T p(s,r,t,m) - d(s,r,t)_2^2\, dt \tag{5}$$

$J_{phase}$ is the exponential phase-based objective function and can be expressed as:

$$J_{phase}(m) = \sum_s\sum_r\int_0^T \left| e^{i\varphi}_{r_syn}(t) - e^{i\varphi}_{r_obs}(t)\right|^2 dt \tag{6}$$

where $e^{i\varphi}_{r_syn}(t)$ and $e^{i\varphi}_{r_obs}(t)$ represent the exponential instantaneous phase information for the r-th receiver from the synthetic and observed data. Bringing **Equation 3** into **Equation 6**, we obtain:

$$J_{phase}(m) = \sum_s\sum_r\int_0^T \left[\frac{s^r(t)}{E_{r_syn}(t)} - \frac{u^r(t)}{E_{r_obs}(t)}\right]^2 dt + \sum_s\sum_r\int_0^T \left[\frac{s^r_H(t)}{E_{r_syn}(t)} - \frac{u^r_H(t)}{E_{r_obs}(t)}\right]^2 dt \tag{7}$$

In **Equation 7**, $S^r(t)$ and $u^r(t)$ represent the $r$-th receiver of the synthetic data and the observed data, respectively, whereas $S^r(t)$ and $u^r_H(t)$ denote the $r$-th receiver of the Hilbert-transformed synthetic data and observed data, respectively. By combining **Equation 5** and **Equation 7**, we construct a weighted instantaneous phase objective function. For the gradient calculation of this objective function, AD is employed: it automatically tracks the forward computation, traverses the computational graph in reverse, and accumulates parameter gradients via the chain rule. This approach eliminates the need to derive adjoint sources for complex objective functions and provides robust computational support for model parameter optimization.

## 3. Results and discussion

### 3.1. Overthrust model tests

This study first employed the Overthrust model through numerical experiments to validate the effectiveness of the proposed method. The true and initial velocity models are shown in Figure 3A and 3B, respectively. The initial velocity model (Figure 3B) was obtained by applying a Gaussian smoothing filter to the true model. The source wavelet has a dominant frequency of 6 Hz. The spatial discretization step is 25 m, and the temporal sampling interval is 4 ms. The model consists of a 10 km horizontal ($X$) extent and a 2.5 km depth ($Z$), with a total recording time of 3 s. The acquisition system included 80 sources spaced at 125 m intervals. For each shot, 401 receivers were deployed along the $X$-direction with a receiver spacing of 25 m. In field seismic data, the low-frequency components are often accompanied by low-frequency noise, making it difficult to obtain low-frequency data with a high signal-to-noise ratio. Moreover, FWI exhibits strong nonlinearity; in the absence of low-frequency components, the inversion is prone to converge to local minima, resulting in unsatisfactory models. To realistically simulate field conditions, frequency components below 3 Hz were removed from the synthetic data in this test.

Figure 4A and 4C show the results of traditional FWI after 40 and 80 iterations, respectively, while Figure 4B and 4D present the results of weighted instantaneous phase inversion at the same number of iterations. The comparative results demonstrate that the weighted instantaneous phase inversion method rapidly constructs a stable overall velocity model that aligns with the true velocity structure in fewer iterations.

Figure 4E shows the inversion result of the traditional FWI. When the initial model significantly differs from the true velocity model, the phase mismatch between the synthetic and observed data leads to cycle-skipping, which prevents the inversion from correctly relating data to model parameters.

Subsequently, we employed the weighted instantaneous

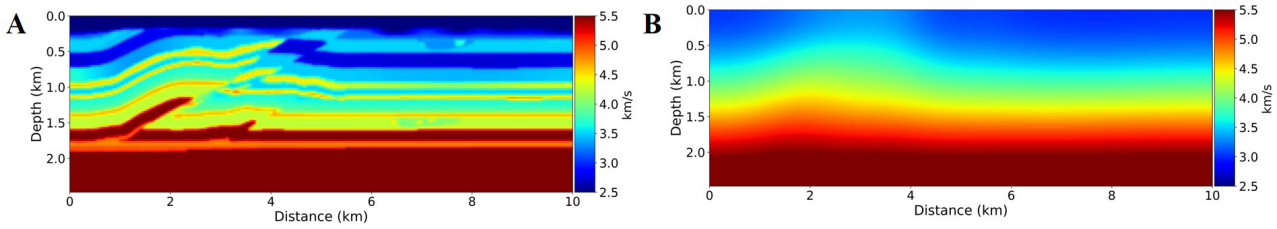



**Figure 3.** Velocity model: (A) True and (B) initial velocity models

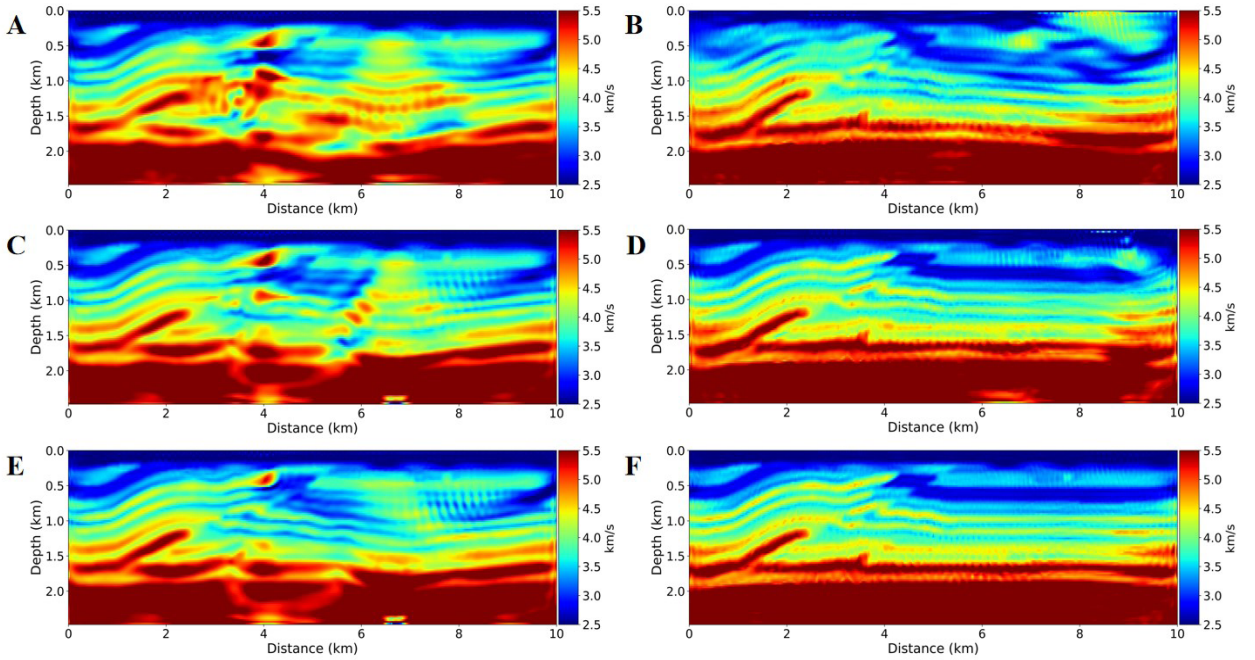



**Figure 4.** Inversion results for the Overthrust model: (A) Traditional full waveform inversion (FWI) after 40 iterations, (B) weighted instantaneous phase inversion after 40 iterations, (C) traditional FWI after 80 iterations, (D) weighted instantaneous phase inversion after 80 iterations, (E) final inversion result of traditional FWI, and (F) final inversion result of weighted instantaneous phase inversion

phase inversion method. During the early stages of inversion, the weight factor was set to 0.2, with the instantaneous phase objective function comprising 80% and the L2 norm objective function 20%. In the later stages, the weighted instantaneous phase had already provided a robust initial model, so the weight factor was increased to 1, with full emphasis on the L2 norm objective function to enhance model resolution. Figure 4F illustrates the result of the weighted instantaneous phase inversion. By comparing Figure 4E and 4F, it is evident that the weighting strategy leverages the complementary strengths of phase and waveform information, maintaining global structural robustness while significantly enhancing the ability to capture details, thereby improving the overall effectiveness and accuracy of the inversion.

Figure 5 presents a comparison of the velocity profile results at a 4-km section. These numerical tests demonstrate that within the framework of AD, the weighted instantaneous phase inversion method effectively mitigates the cycle-skipping issue in FWI while enhancing the accuracy of velocity inversion, achieving robust inversion in complex geological environments.

### 3.2. Marmousi model tests

To validate the applicability of the weighted instantaneous phase inversion method based on AD in more complex models, we extended our tests from the existing Overthrust model to the Marmousi model. The Marmousi model, with its intricate layered structures and diverse velocity field characteristics, poses a greater challenge to the inversion process. The true model used for inversion is shown in Figure 6A, whereas Figure 6B illustrates the linear initial model employed in the inversion. Compared to the initial model derived from smoothing the true model, the linear

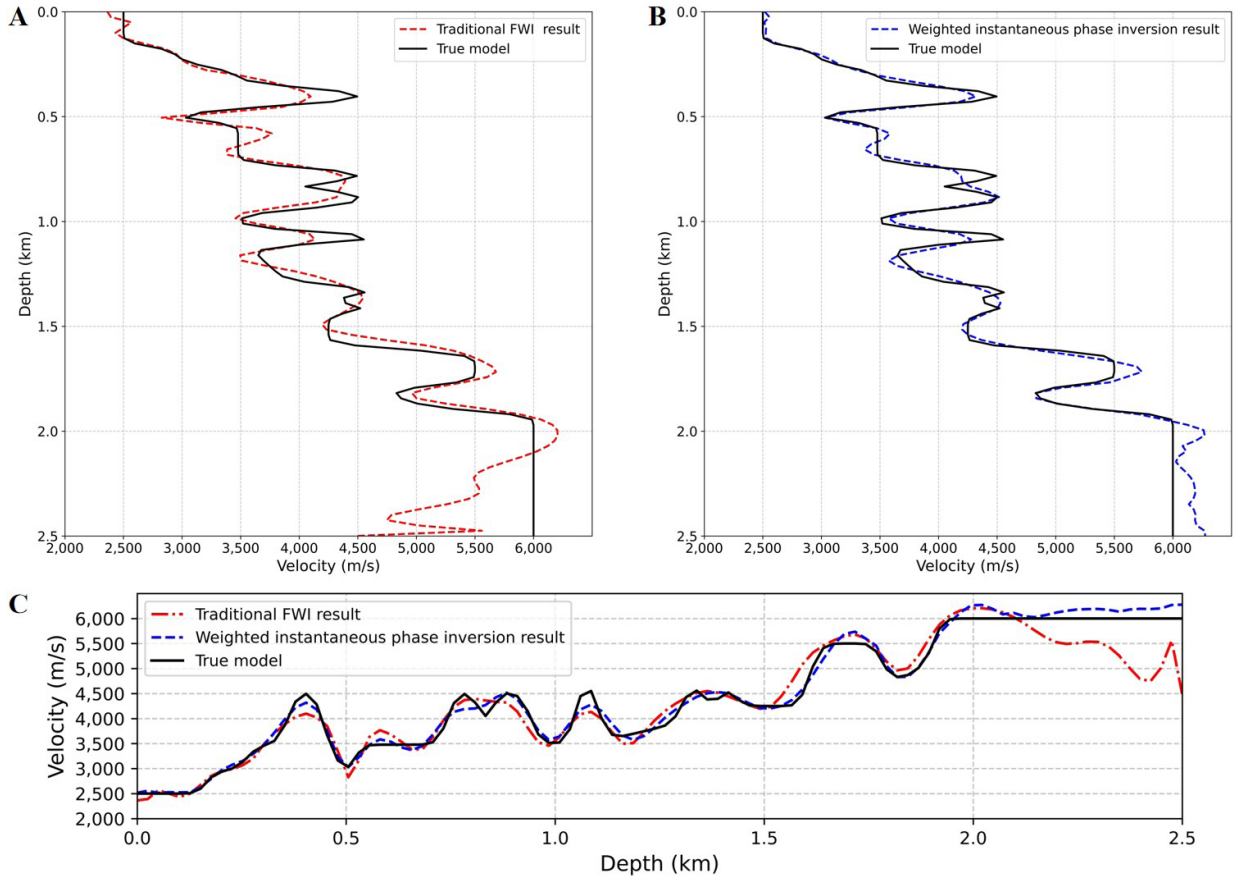



**Figure 5.** Velocity profiles extracted at a 4-km section: (A) Result from traditional full waveform inversion (FWI), (B) result from weighted instantaneous phase inversion, and (C) comparison of velocity profiles

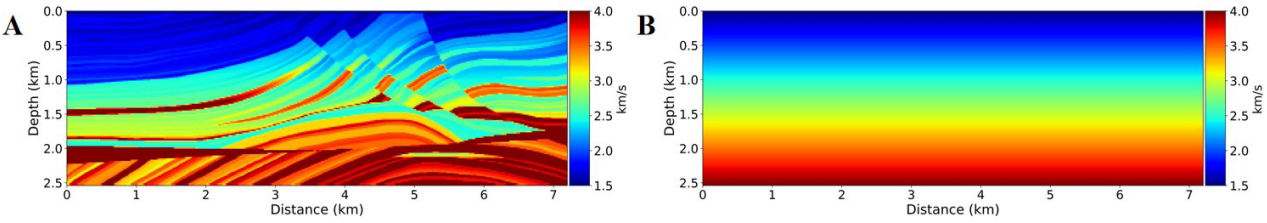



**Figure 6.** Inverted model: (A) True and (B) initial model

model lacks large-scale structures, making the objective function more susceptible to convergence to local minima.

Figure 7A and 7C present the results of the traditional FWI after 40 and 80 iterations, respectively, while Figure 7B and 7D show the results of the weighted instantaneous phase inversion method for the same number of iterations. The comparison demonstrates that, even for the complex Marmousi model, the weighted instantaneous phase inversion method rapidly produces a stable and accurate overall velocity model within relatively few iterations. Figure 7E and 7F present the final inversion results. By comparing Figure 7E and 7F, it is evident that the weighted instantaneous phase inversion method achieves satisfactory results even in the presence of complex layered structures. Particularly at a depth of 1.5 km on the right side, the method successfully recovers the precise velocity field, avoiding the inversion errors and blurred boundaries observed in Figure 7E.

Figure 8 compares the velocity profiles extracted at a distance of 5 km. Figure 8A shows that the traditional FWI exhibits noticeable deviations in the middle and deeper layers. In contrast, Figure 8B illustrates the velocity profile

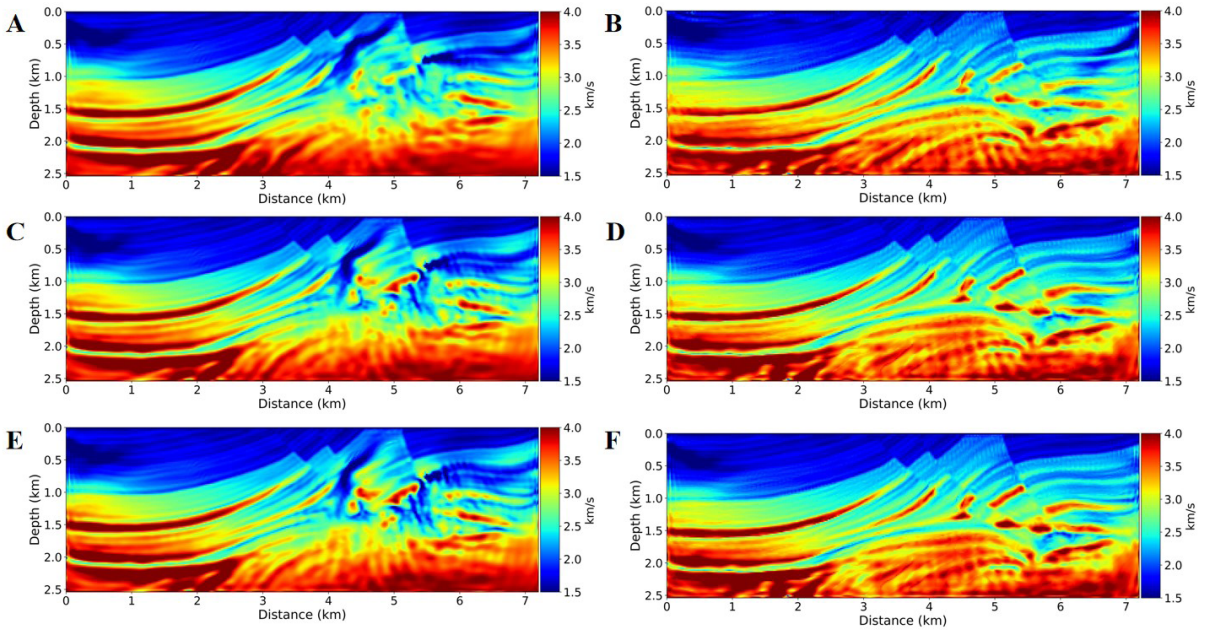



**Figure 7.** Inversion results for the Marmousi model: (A) Traditional full waveform inversion (FWI) after 40 iterations, (B) weighted instantaneous phase inversion after 40 iterations, (C) traditional FWI after 80 iterations, (D) weighted instantaneous phase inversion after 80 iterations, (E) final inversion result of traditional FWI, and (F) final inversion result of weighted instantaneous phase inversion

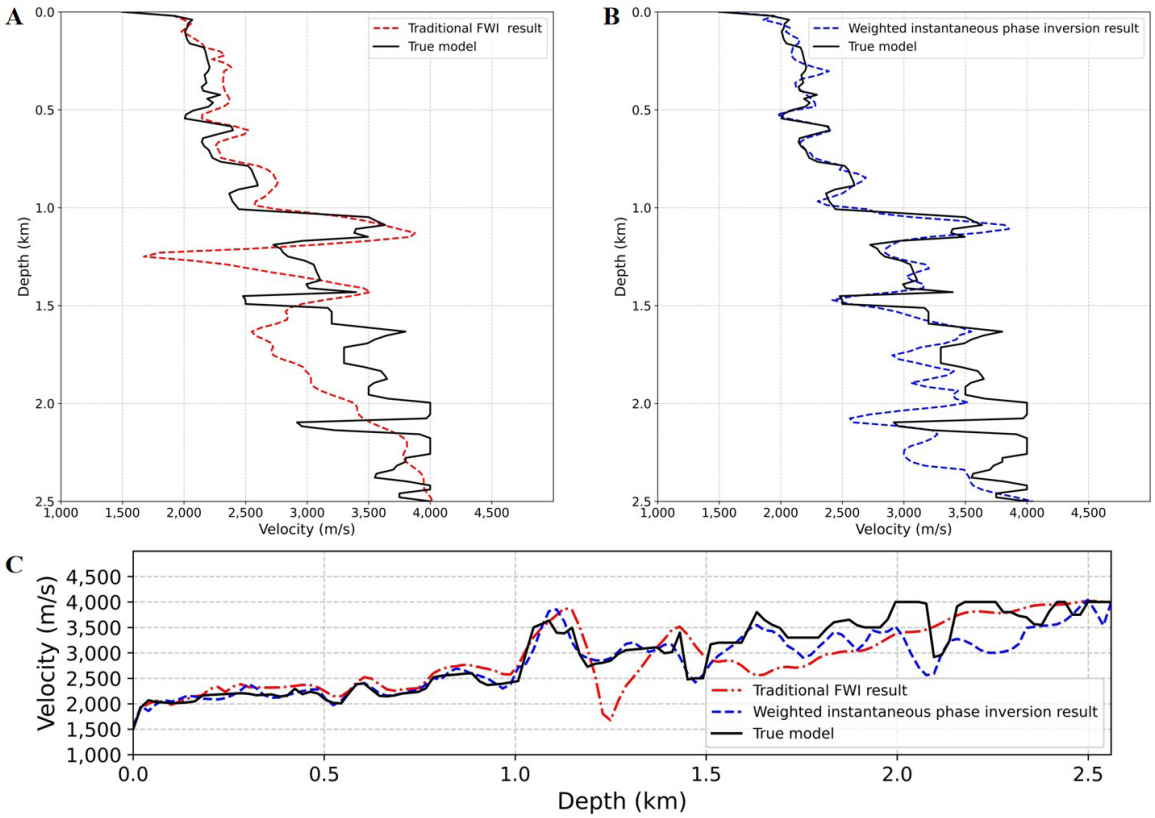



**Figure 8.** Velocity profiles extracted at a 5-km section: (A) Result from traditional full waveform inversion (FWI), (B) result from weighted instantaneous phase inversion, and (C) comparison of velocity profiles

using the weighted instantaneous phase objective function, which maintains higher accuracy in both shallow and deep layers compared to the traditional FWI, demonstrating a closer match with the true model. Figure 8C further shows that the weighted instantaneous phase objective function better characterizes the velocity model. The method effectively alleviates the cycle-skipping issue and enhances the stability and resolution of the inversion process.

To visually compare the inversion results with the true model, Figure 9 illustrates the absolute error distributions for the two objective functions. Figure 9A shows the absolute error between the traditional FWI results and the true model, with large residual errors in the middle and lower depths on the right side. Conversely, Figure 9B presents the absolute error distribution obtained using the weighted instantaneous phase objective function, demonstrating a notable reduction in error in the same region compared to Figure 9A.

Testing with both the Overthrust and the Marmousi models demonstrates that employing a weighted instantaneous phase inversion strategy, facilitated by AD, significantly improves the robustness and resolution of the inverted velocity model. It effectively mitigates the cycle-skipping problem typically arising from the absence of low-frequency information in FWI, resulting in more accurate inversion outcomes.

To quantitatively compare the two inversion methods, Table 1 presents the root mean square error and mean absolute error of the inversion results relative to the true model. As shown in Table 1, the weighted instantaneous

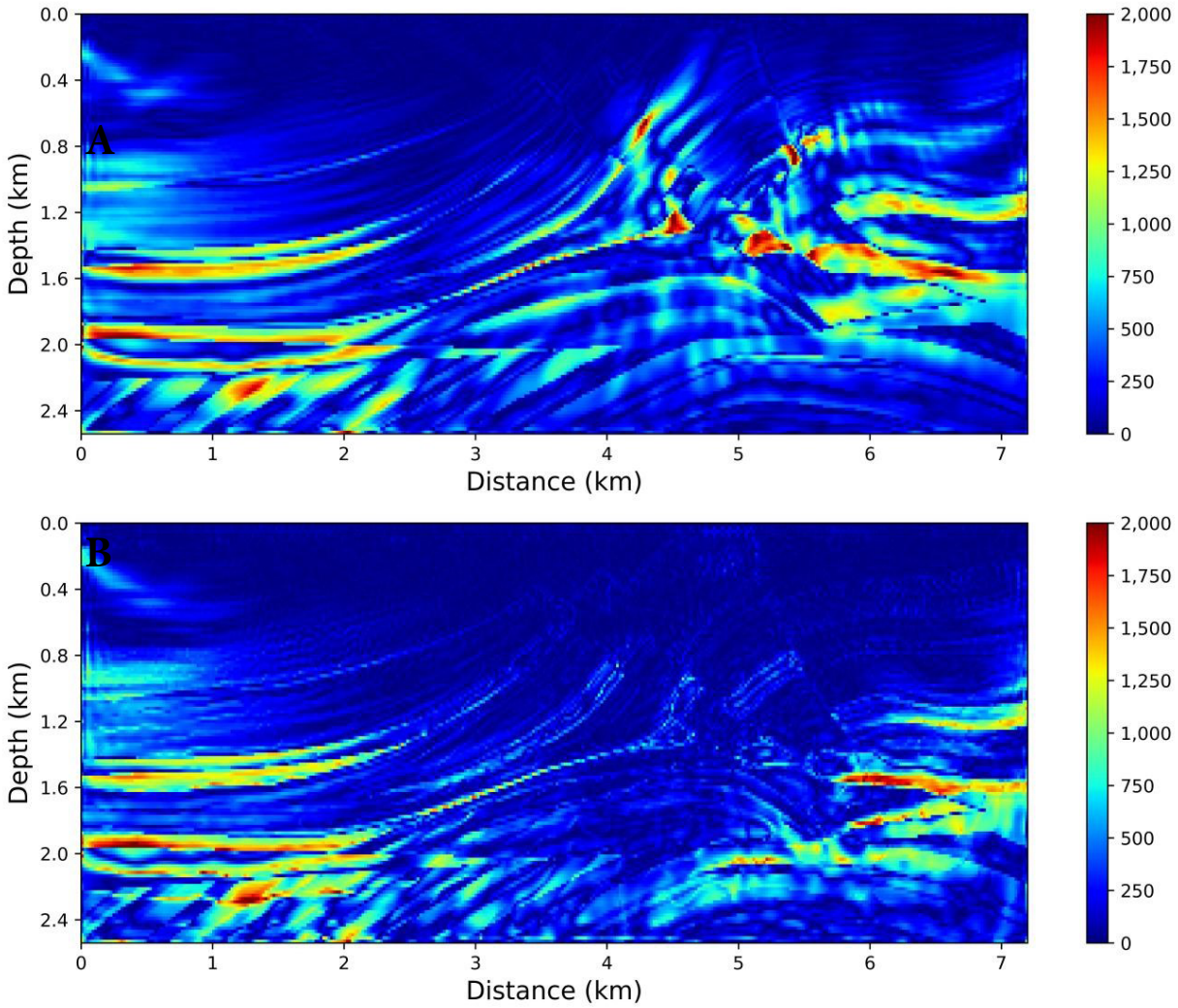



**Figure 9.** Comparison of absolute errors between the inversion results and the true model. (A) traditional full waveform inversion; (B) weighted instantaneous phase inversion.

**Table 1. Root mean square error and mean absolute error of inversion results for the Overthrust and Marmousi models compared to the true model**

| Model | Method | Weighted instantaneous phase inversion | Traditional FWI |
|---|---|---|---|
| Overthrust | RMSE | 214.45 | 423.76 |
| | MAE | 144.09 | 307.15 |
| Marmousi | RMSE | 413.42 | 485.17 |
| | MAE | 253.91 | 328.46 |

Abbreviations: FWI: Full waveform inversion; MAE: Mean absolute error; RMSE: Root mean square error.

phase inversion produces smaller errors and achieves results closer to the true model.

### 3.3 Noise testing

Seismic data collected in the field often include noise, which can significantly affect FWI results. To evaluate the noise resistance of the weighted instantaneous phase inversion method, Gaussian noise was added to the observed data. Figure 10A shows the noise-free seismic records from the Overthrust model, while Figure 10B shows the noisy seismic records obtained by adding Gaussian noise. These noisy observations were then used to perform FWI.

Figure 11 shows the inversion results using the weighted instantaneous phase inversion method. Figure 11A and 11B present the inversion results without and with noise, respectively. The comparison reveals the method's strong noise resistance. In Figure 11B, despite the noise interference, key geological structures remain clearly identifiable. In contrast, Figure 11A demonstrates higher resolution and inversion accuracy in the absence of noise. Overall, the weighted instantaneous phase inversion method maintains the core quality of inversion results under noisy conditions, confirming its robustness and reliability for real exploration scenarios.

Gaussian noise was also added to the seismic records of the Marmousi model. Figure 12 compares the seismic data before and after noise addition. Figure 13 shows the inversion results for the Marmousi model using the weighted instantaneous phase method. Comparing the noisy inversion results (Figure 13B) with the noise-free results (Figure 13A) shows that, even for the complex

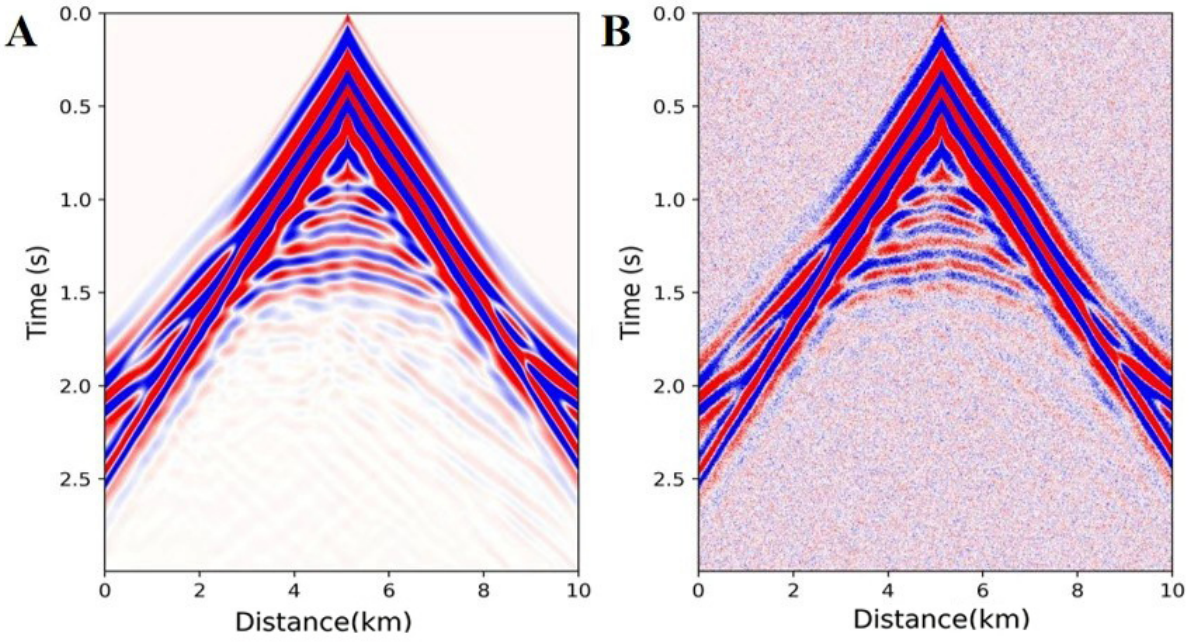



**Figure 10.** Seismic records from the Overthrust model: (A) Observed data without noise and (B) observed data with added Gaussian noise

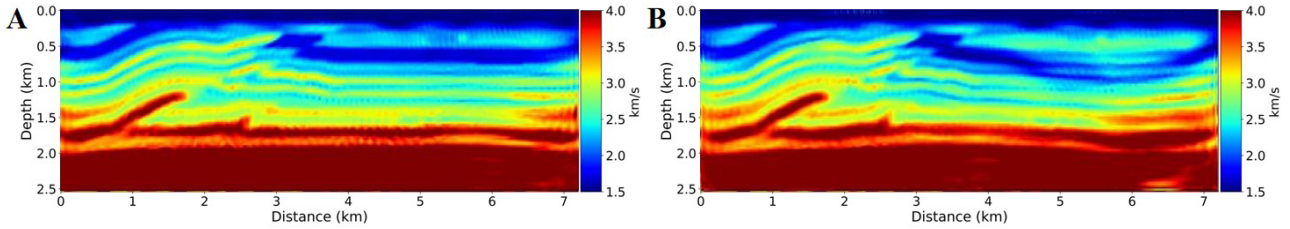



**Figure 11.** Weighted instantaneous phase inversion results for the Overthrust model: (A) Inversion result without noise and (B) inversion result with added Gaussian noise

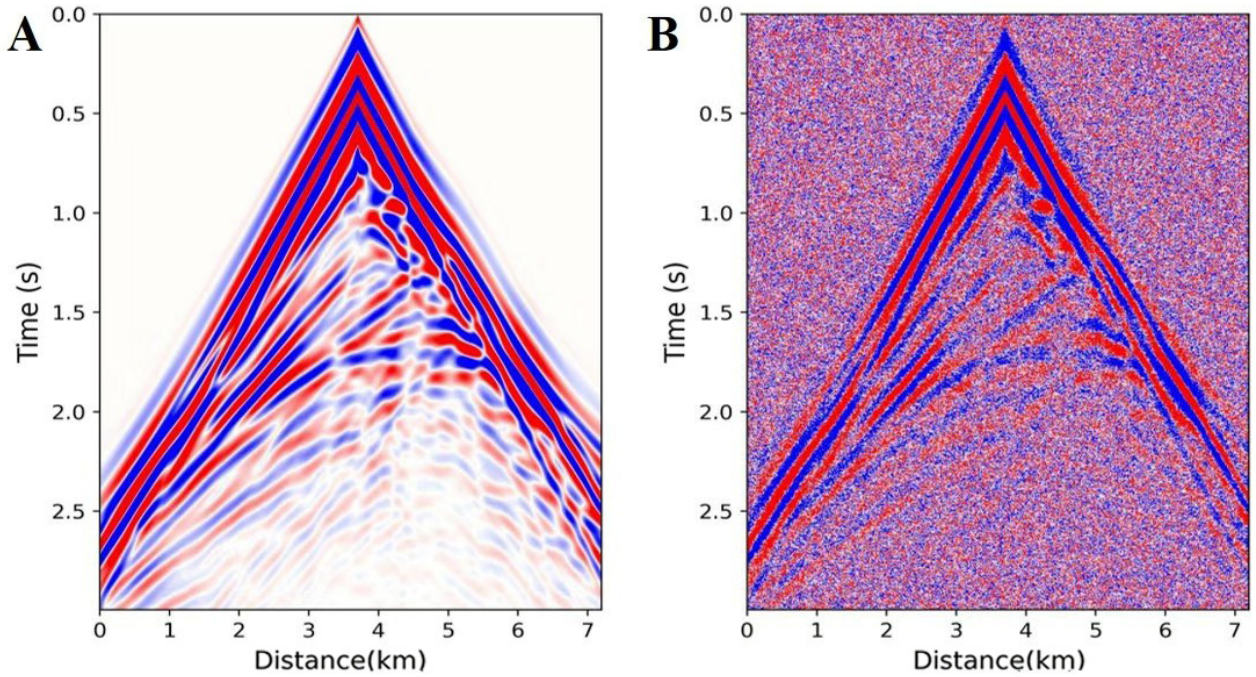



**Figure 12.** Seismic records from the Marmousi model: (A) Observed data without noise and (B) observed data with added Gaussian noise

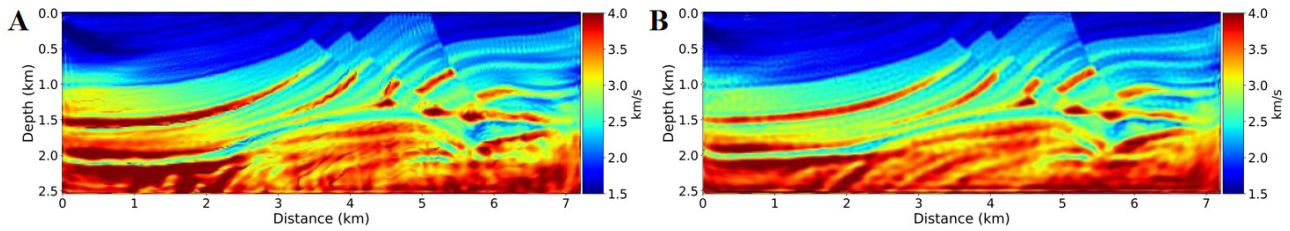



**Figure 13.** Weighted instantaneous phase inversion results for the Marmousi model: (A) Inversion result without noise and (B) inversion result with added Gaussian noise

Marmousi model, the method demonstrates strong noise resistance, successfully recovering an accurate subsurface velocity model.

## 4. Conclusion

This study introduces an AD-based weighted instantaneous phase inversion method to address the cycle-skipping problem in FWI. The relationship between the instantaneous phase information of seismic signals and the subsurface velocity structure is more linear, while waveform information facilitates high-resolution results and maintains algorithmic stability. Therefore, within the AD-based waveform inversion framework, we combine instantaneous phase information with waveform information to construct a weighted instantaneous phase objective function, achieving a balanced trade-off between inversion robustness and resolution. To utilize instantaneous phase information effectively, we avoided phase wrapping issues using exponential phase information. Numerical tests indicate that the early incorporation of instantaneous phase information provides a reasonable initial velocity model, mitigates the cycle-skipping problem, and significantly enhances overall inversion performance.

## Acknowledgments

None.

## Funding

This research was funded by the National Natural Science Foundation of China (Grant No. 42104116).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Mengzi Wang, Yong Hu
*Formal analysis:* Chunyuan Shi, Zhihan Zhang
*Investigation:* Mengzi Wang, Yong Hu, Zhihan Zhang
*Methodology:* Mengzi Wang, Yong Hu
*Software:* Mengzi Wang, Zhihan Zhang
*Supervision:* Mengzi Wang, Yong Hu
*Writing–original draft:* Mengzi Wang, Chunyuan Shi
*Writing–review & editing:* Xiaoyuan Liao

## Availability of data

Data are available from the corresponding author upon reasonable request.

## References

1. Lailly P. The seismic inverse problem as a sequence of before stack migrations. In: Bednar JB, Robinson E, Weglein A, eds. *Conference on Inverse Scattering—Theory and Application*. Philadelphia, PA: SIAM; 1983:206-220.
2. Tarantola A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*. 1984;49(8):1259-1266.

   doi: 10.1190/1.1441754
3. Fichtner A, Bunge HP, Igel H. The adjoint method in seismology: I. Theory. *Phys Earth Planet Inter*. 2006;157(1-2):86-104.

   doi: 10.1016/j.pepi.2006.03.016
4. Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics. *Geophysics*. 2009;74(6):WCC1-WCC26.

   doi: 10.1190/1.3238367
5. Plessix RE. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys J Int*. 2006;167(2):495-503.

   doi: 10.1111/j.1365-246X.2006.02978.x
6. Luo JR, Wang BF. Initial model building in time domain elastic full waveform inversion using the instantaneous phase information. *Prog Geophys*. 2018;33(6):2435-2440. [In Chinese].

   doi: 10.6038/pg2018CC0215
7. Sun SY, Hu GH, He BH, Du ZY, Xu WC. Reflection waveform inversion and its application to onshore seismic data. *Prog Geophys*. 2021;36(6):2566-2572. [In Chinese].

   doi: 10.6038/pg2021EE0462
8. Li ZC, Wang ZY, Huang JP, Cui C. Multi-scale full waveform inversion based on gradient decomposition in wavenumber domain. *Chin J Geophys*. 2022;65(7):2693-2703. [In Chinese].

   doi: 10.6038/cjg2022P0508
9. Wang QQ, Song P, Hua QF, *et al*. Full waveform inversion based on Adam algorithm with optimized parameters. *Chin J Geophys*. 2023;66(11):4654-4663. [In Chinese].

   doi: 10.6038/cjg2022Q0879
10. Li JS, Liu WK, Liang YX, Wu JY, Li ZG. Time-domain full waveform inversion based on high-order amplitude information. *Prog Geophys*. [In Chinese]. 2023;38(4):1603-1609.

    doi: 10.6038/pg2023GG0129
11. Kang PF, Hu Y, Liu RD, *et al*. Local-scale frequency-wavenumber domain phase inversion. *Prog Geophys*. 2025;40(1):155-165. [In Chinese].

    doi: 10.6038/pg2025II0034
12. Shah NK, Warner MR, Washbourne JK, Guasch L, Umpleby AP. A phase-unwrapped solution for overcoming a poor starting model in full wavefield inversion. In: *Proceedings*

*of the 74th EAGE Conference and Exhibition Incorporating EUROPEC 2012;* Copenhagen, Denmark: Jun 4-7, 2012; European Association of Geoscientists & Engineers; 2012:cp-293-00460.

doi: 10.3997/2214-4609.20148455

13. Choi Y, Alkhalifah T. Unwrapped phase inversion with an exponential damping. *Geophysics*. 2015;80(5):R251-R264.

doi: 10.1190/geo2014-0498.1

14. Luo JR, Wu RS, Gao FC. Time-domain full waveform inversion using instantaneous phase information with damping. *J Geophys Eng*. 2018;15(3):1032-1041.

doi: 10.1088/1742-2140/aaa984

15. Hu Y, Wu RS, Han LG, Zhang P. Joint multiscale direct envelope inversion of phase and amplitude in the time–frequency domain. *IEEE Trans Geosci Remote Sens*. 2019;57(7):5108-5120.

doi: 10.1109/TGRS.2019.2896936

16. Zhu HJ, Fomel S. Building good starting models for full-waveform inversion using adaptive matching filtering misfit. *Geophysics*. 2016;81(5):U61-U72.

doi: 10.1190/geo2015-0596.1

17. Warner M, Guasch L. Adaptive waveform inversion: Theory. *Geophysics*. 2016;81(6):R429-R445.

doi: 10.1190/geo2015-0387.1

18. Hu Y, Han LG, Yu JL, Chen RD. Time-frequency domain multi-scale full waveform inversion based on adaptive non-stationary phase correction. *Chin J Geophys*. 2018;61(7):2969-2988. [In Chinese].

doi: 10.6038/cjg2018L0421

19. Liu JG, Xu KL. Automatic differentiation and its applications in physics simulation. *Acta Phys Sin*. 2021;70(14):149402. [In Chinese].

doi: 10.7498/aps.70.20210813

20. Ketkar N, Moolayil J. Automatic Differentiation in Deep Learning. In: *Deep Learning with Python*. Berkeley, CA: Apress; 2021:133-145.

doi: 10.1007/978-1-4842-5364-9_4

21. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.

doi: 10.1038/nature14539

22. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press; 2016.

23. Nolan JF. *Analytical differentiation on a digital computer* [master's thesis]. Cambridge, MA: Massachusetts Institute of Technology, Department of Mathematics; 1953. https://hdl.handle.net/1721.1/12297

24. Speelpenning B. *Compiling Fast Partial Derivatives of Functions Given by Algorithms* [PhD thesis]. Urbana, IL: University of Illinois at Urbana-Champaign, Department of Computer Science; 1980. Report No. UIUC DS-R-80-1002.

doi: 10.2172/5254402

25. Kasim MF, Vinko SM. ξ-torch: differentiable scientific computing library. *arXiv*. Preprint posted online October 6, 2020.

doi: 10.48550/arXiv.2010.01921

26. Sambridge M, Rickwood P, Rawlinson N, Sommacal S. Automatic differentiation in geophysical inverse problems. *Geophys J Int*. 2007;170(1):1-8.

doi: 10.1111/j.1365-246X.2007.03400.x

27. Tan L, Brytik V, Baumstein A, Hinkley D. Verification of gradient and hessian computation for full wavefield inversion using automatic differentiation. In: Levin S, ed. SEG Technical Program Expanded Abstracts 2010; SEG International Exposition and 80th Annual Meeting; 17–22 October 2010; Denver, CO, USA. Society of Exploration Geophysicists; 2010:2762-2766.

doi: 10.1190/1.3513417

28. Liao W. An accurate and efficient algorithm for parameter estimation of 2D acoustic wave equation. *Int J Appl Phys Math*. 2011;1(2):96-100.

doi: 10.7763/IJAPM.2011.V1.19

29. Richardson A. Seismic full-waveform inversion using deep learning tools and techniques. *arXiv*. Preprint posted online January 22, 2018.

doi: 10.48550/arXiv.1801.07232

30. Cao D, Liao W. A computational method for full waveform inversion of crosswell seismic data using automatic differentiation. *Comput Phys Commun*. 2015;188:47-58.

doi: 10.1016/j.cpc.2014.11.002

31. Zhu W, Xu K, Darve E. A general approach to seismic inversion with automatic differentiation. *Comput Geosci*. 2021;151:104751.

doi: 10.1016/j.cageo.2021.104751

32. Song C, Wang Y, Richardson A, Liu C. Weighted envelope correlation-based waveform inversion using automatic differentiation. *IEEE Trans Geosci Remote Sens*. 2023;61:1-11.

doi: 10.1109/TGRS.2023.3300127

33. Liu F, Li H, Zou G, Li J. Automatic differentiation-based full waveform inversion with flexible workflows. *J Geophys Res Mach Learn Comput*. 2025;2(1):e2024JH000542.

doi: 10.1029/2024JH000542

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Direct envelope-based reflection full waveform inversion for shallow marine seismic data

**Kyoungmin Lim[1], Wookeen Chung[1], Jungkyun Shin[2], and Jiho Ha[2]***

[1]Department of Energy and Resources Engineering, National Korea Maritime & Ocean University, Busan, Republic of Korea

[2]Pohang Research Branch, Korea Institute of Geoscience and Mineral Resources, Pohang, Republic of Korea

(This article belongs to the *Special Issue: Full Waveform Inversion Methods and Applications for Seismic Data in Complex Media*)

## Abstract

Accurate velocity models that include long-wavelength components are required for precise subsurface structure imaging and estimation of geophysical properties. To successfully build the long-wavelength velocity using full waveform inversion (FWI), sufficient offset and low-frequency components are necessary. However, due to the limited acquisition conditions, it is difficult to operate long offset and low-frequency sources in coastal shallow marine. Short offset makes the data dominated by reflections, and FWI intensively updates the surface boundaries. Although reflection full waveform inversion (RFWI) has been proposed to reconstruct long-wavelength velocity models using reflection data, it suffers from cycle skipping when low-frequency component is insufficient. To overcome these limitations, we propose a direct envelope-based RFWI (DE-RFWI) that incorporates the direct envelope into the RFWI. By employing envelope-based energy information from reflection data, DE-RFWI facilitates the reconstruction of long-wavelength velocity. The proposed method employs the Hilbert transform-based implicit gradient decomposition technique to address additional computational cost. To verify the proposed method, DE-RFWI was applied to synthetic test with a shallow marine condition and field data acquired in Yeongil Bay, South Korea. The inversion results for field data were evaluated by analyzing arrival-time alignment in envelope domain. The results demonstrate that DE-RFWI can reliably reconstruct long-wavelength velocity models in shallow marine seismic data and improve reflector continuity and resolution in the reverse time migration imaging.

***Keywords:*** Long-wavelength velocity model; Shallow marine seismic; Short offset; Reflection full waveform inversion; Direct envelope

***Corresponding author:**
Jiho Ha
(jihoha@kigam.re.kr)

**Citation:** Lim K, Chung W, Shin J, Ha J. Direct envelope-based reflection full waveform inversion for shallow marine seismic data. *J Seismic Explor*. 2026;35(2):026030005. doi: 10.36922/JSE026030005

**Received:** January 14, 2026

**Revised:** February 25, 2026

**Accepted:** March 9, 2026

**Published online:** March 27, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Recently, increasing interest in coastal regions, such as the offshore wind farms, stability assessment of marine structures, submarine geological monitoring, and geo-hazard detection, has highlighted the growing need for accurate characterization of

subsurface elastic properties and velocity structures. Multi-beam echo sounder (MBES) and sub-bottom profiler (SBP) are survey systems widely used in coastal region. Since MBES and SBP operate at high frequencies, these systems provide only information on the seafloor and the shallow subsurface.[1-3] In contrast, marine seismic surveys that use air-gun and sparker cover a frequency range within the tens-of-hertz to kilohertz band, providing information on deeper subsurface structures. However, conventional seismic survey for reservoir exploration requires long streamer arrays and large vessels, limiting their applicability in coastal areas where the water depth is shallow and various fishing activities are conducted. To overcome these challenges, ultra-high resolution (UHR) seismic survey has been adopted.[4,5] UHR surveys can be operated from small vessels, reduce the group interval, and utilize higher source frequencies, enabling the acquisition of high-resolution seismic data.

Seismic signals acquired in coastal shallow marine are dominantly recorded as reflections because the short source-receiver offset restrict the propagation paths. Although this characteristic enhances the resolution of seismic sections, it suffers from a lack of wide-aperture. Wide-aperture signals are crucial for velocity model building and have been utilized to reconstruct long-wavelength components of the subsurface structure.[6-9]

Full waveform inversion (FWI) is a technique that estimates high-resolution subsurface properties by minimizing the misfit between modeled and observed waveforms. Under ideal conditions, FWI can update both subsurface reflectors and the background velocity, which are controlled by the migration and tomography components, respectively.[10] However, because FWI relies on wide-aperture information to build a long-wavelength velocity model,[11,12] it becomes challenging to obtain a reliable long-wavelength model on reflection dominant data.

In reflection dominant data, the migration component is much stronger than the tomography component, resulting in boundary-limited updates. Reflection full waveform inversion (RFWI) was proposed[13] to recover long-wavelength velocity model with reflection data. Table 1 summarizes representative prior studies on RFWI, classified according to their primary research focus. The Freq. and Max. offset columns describe the data specifications used in the numerical tests. The frequency not given as a range denotes the dominant frequency. Early studies recognized that gradient derived from reflection data is migration dominant, and gradient decomposition strategies were introduced to enhance the tomographic component.[13-17] More recently, RFWI has been extended to elastic and anisotropic media,[18-20] and efforts have been made to improve its computational efficiency.[21-23]

To enhance the tomographic component in reflection data, Xu *et al.*[13] decomposed the gradient into migration and tomography component by separating the incident and scattered wavefields using a Born modeling. Born modeling-based approaches have been widely used because the first-order scattered wavefield can be explicitly isolated.[24-26] However, such approaches are based on the Born approximation to simulate the first order scattering, and it is difficult to account for higher-order scattering effects.[27] Wavefield decomposition methods separate the source and receiver wavefields into up/down-going according to their propagation directions, from which the migration and tomography components can be constructed. This strategy can effectively separate each component in shallow seismic conditions where wave propagation is relatively simple and clear. Therefore, in this study, we adopted a wavefield decomposition method to separate the migration and tomography components.

RFWI can reconstruct long-wavelength velocity models using reflection data. However, because conventional RFWI employs the waveform matching–based objective function, it becomes vulnerable to cycle skipping and local minima.[25] To mitigate this issue, the correlation-based objective function that emphasizes kinematic similarity[28] and envelope-based objective function that suppresses phase dependence[29,30] were proposed. However, conventional envelope inversion calculates the waveform-based Fréchet derivative through the chain rule, and the gradient can be distorted in the reflection data. To overcome these issues, Wu and Chen[31] proposed direct envelope inversion (DEI), which can stably estimate long-wavelength velocity under strong scattering conditions by directly deriving the envelope Fréchet derivative. DEI has been successfully applied to reconstruct large-scale velocity anomalies with strong velocity contrasts.[32-34]

Chen *et al.*[29] demonstrated that multi-scale DEI prior to multi-scale RFWI leads to reliable long-wavelength velocity model and improves subsalt imaging. They applied wavefield decomposition to DEI, in which the migration and tomography components were alternately utilized to update the short- and long-wavelength components, respectively. Although their approach achieved successful results in synthetic experiments, the decomposition method in $f$-$k$ domain requires storing both source and receiver wavefields at every time steps, resulting in large memory consumption. This limitation becomes more critical in shallow marine seismic data, where high-frequency data and dense temporal samples are required. Therefore, in this study, we adopt an efficient wavefield decomposition

technique based on the Hilbert transform proposed by Lian *et al.*[35] This method can implicitly decompose the wavefields in time-space domain by constructing analytic signals using the Hilbert transform to separate up/down-going components. Since each component can be implicitly extracted in time-space domain, 2D FFT is unnecessary to transform $f$ - $k$ domain and storing the full wavefields can be avoided.

Previous RFWI studies have generally been conducted using datasets with offsets of several kilometers and frequency bands below several tens of hertz. Although short-offset targeted study has been reported,[36] it still relied on offset ranges exceeding 1 km. In this study, we aim to reconstruct a long-wavelength velocity model from shallow marine seismic data. Accordingly, the RFWI is proposed to recover long-wavelength velocity information from reflection dominant data. Shallow seismic data are characterized by severely limited offsets and the difficulty of utilizing low-frequency sources, resulting in wavefields dominated by strong boundary reflections. These characteristics are similar to the strong-scattering

**Table 1. Representative prior studies on RFWI**

| Category | Year | Authors | Remarks | Freq. | Max. offset |
|---|---|---|---|---|---|
| Gradient decomposition | 2012 | Xu *et al.*[13] | Proposed the RFWI to recover the long-wavelength velocity model from reflected waves | ~8 Hz | 8 km |
| | 2013 | Tang *et al.*[14] | Decomposed the gradient into tomography and migration components and emphasized the tomography component using weight factor | 25 Hz | - |
| | 2014 | Alkhalifah[15] | Proposed the scattering-angle-guided filter for gradient decomposition | - | - |
| | 2016 | Wang *et al.*[16] | Proposed the alternating inversion of long- and short-wavelength components using wavefield decomposition | 5–40 Hz | 1.2 km |
| | 2018 | Yao *et al.*[17] | Proposed the plane-wave angle filter to separate migration and tomography components | ~24 Hz | 10 km |
| Objective function | 2015 | Chi *et al.*[28] | Introduced the correlation-based objective function for RFWI | - | 3.875 km |
| | 2015 | Wu & Alkhalifah[37] | Proposed simultaneous inversion of background and perturbation using modified objective function | - | - |
| | 2018 | Chen *et al.*[29] | Proposed the multi-scale envelope inversion using WAE to construct start model of multi-scale RFWI | 4–40 Hz | 7 km |
| | 2024 | Wang *et al.*[30] | Proposed the envelope-normalized objective function for RFWI | - | 4 km |
| | **2026** | **Lim *et al.* (this paper)** | **Proposed the direct envelope-based RFWI to reconstruct long-wavelength velocity model on shallow marine seismic data** | **~200 Hz** | **90 m** |
| Elastic media | 2017 | Guo & Alkhalifah[18] | Introduced the elastic RFWI to update P and S long-wavelength velocity using equivalent stress source | 2–8 Hz | 6 km |
| Anisotropy | 2020 | Li & Alkhalifah[19] | Proposed the multi-parameter RFWI for acoustic VTI media | 2–8 Hz | 3 km |
| | 2023 | Wu *et al.*[20] | Introduced the Gauss–Newton multi-parameter RFWI for VTI media. | 5–25 Hz | 4 km |
| Computational efficiency | 2021 | Song & Alkhalifah[21] | Introduced EWI using subsurface secondary source into RFWI to improve computational efficiency | 3–10 Hz | 3 km |
| | 2021 | Wang *et al.*[22] | Proposed the matrix-free Gauss–Newton RFWI to improve convergence speed | 6–25 Hz | 8 km |
| | 2022 | Xu *et al.*[23] | Proposed the second-order Gauss–Newton optimization to RFWI to improve convergence speed and resolution | - | 4 km |

Abbreviations: EWI: Efficient wavefield inversion; RFWI: Reflection full waveform inversion; VTI: Vertical transversely isotropic; WAE: Window-averaged envelope.

conditions assumed in DEI.

Based on these similarities, we propose a direct envelope–based RFWI (DE-RFWI) strategy that incorporates DEI into the RFWI to reconstruct long-wavelength velocity models from shallow marine seismic data. RFWI is adopted to overcome the insufficient wide-aperture information due to limited offset, and DEI is employed to mitigate low-frequency component deficiency and phase sensitivity. The effectiveness of the DE-RFWI approach is demonstrated through numerical examples, and its applicability is further validated using field data.

## 2. Methodology

### 2.1 Full waveform inversion (FWI)

Full waveform inversion (FWI) reconstructs high-resolution subsurface properties by minimizing the misfit between modeled and observed data. Because FWI exploits both the amplitude and phase of the entire waveform, it is capable of estimating velocity models. FWI is formulated as an optimization problem that minimizes the following misfit function:[38]

$$J(\mathbf{m}) = \frac{1}{2}\sum_{s}\sum_{r}\int\left[u(t,m,r;s) - d(t,r;s)\right]^2 dt \tag{1}$$

where $u$ is modeled data and $d$ is observed data with $s$ and $r$ indicating the locations of source and receiver. The gradient can be calculated by taking a partial derivative with respect to the model parameters $\mathbf{m}$:

$$\mathrm{g} = \frac{\partial J(\mathbf{m})}{\partial \mathbf{m}} = \sum_{s}\sum_{r}\int \frac{\partial u(t,r;s)}{\partial \mathbf{m}} \cdot R(t,r;s)\, dt \tag{2}$$

where g is gradient, $(\partial u / \partial \mathbf{m})$ is the partial derivative wavefield, and $R = u - d$ is residual between modeled data and observed data, and $\cdot$ indicates the dot product operator. The partial derivative wavefield can be expressed by using the Green's function $\mathbf{G}$ and virtual source wavefield $\mathbf{v}$:[39]

$$\frac{\partial u}{\partial \mathbf{m}} = \mathbf{G}(t) * \mathbf{v}(t) = \mathbf{v}(t) * \mathbf{G}(t) \tag{3}$$

where * is a convolution operator in time. By substituting **Equation 3** into **Equation 2**, the gradient can be written as:

$$\begin{aligned} g &= \sum_{s}\sum_{r}\left(\mathbf{v}(t) * \mathbf{G}(t)\right) \cdot R(t,r;s) \\ &= \sum_{s}\sum_{r}\mathbf{v}(t) * \left(\mathbf{G}(t) * R(T_{max} - t, r; s)\right) \\ &= \sum_{s}\sum_{r}\mathbf{v}(t) * \mathbf{B}(t) \end{aligned} \tag{4}$$

where $\mathbf{B}$ is the back-propagated wavefield calculated by propagating the time-reverse residual wavefield.

### 2.2. Reflection full waveform inversion (RFWI)

When sufficient offset and broad frequency bandwidth are available, FWI can effectively recover not only reflector boundaries through the migration component but also the background velocity through the tomography component.[10] However, in limited offset conditions, FWI performs updates primarily along boundaries. In the absence of wide-aperture energy, most of the recorded signals consist of reflections. The reflection kernel in FWI appears as a mixture of the oval-shaped migration and the rabbit-ear-shaped tomography component (Figure 1A). Because the migration component is much stronger than tomography component, FWI applied to reflection data fails to reconstruct the background velocity.

To overcome this limitation, Xu *et al.*[13] proposed RFWI, which can stably update the long-wavelength velocity model by selectively utilizing tomography components. By separating migration and tomography components from the full reflection kernel, each component can be independently emphasized. Figure 1B shows the highlights of the oval-shaped migration component, whereas Figure 1C presents the rabbit-ear-shaped tomography component with weakened migration component. The essential concept of RFWI is to decompose the tomography component, which contains long-wavelength information.[27]

A wavefield can be separated into up/down-going according to the propagation direction, which can be explicitly decomposed in the frequency-wavenumber domain by performing a two-dimensional fast Fourier transform (2D FFT) as follows:[40]

$$u(\omega, k_z) = \begin{cases} u_u(\omega, k_z), if\ \omega k_z < 0 \\ u_d(\omega, k_z), if\ \omega k_z \geq 0 \end{cases} \tag{5}$$

where $u_u$ is the up-going wavefield and $u_d$ is the down-going wavefield. From **Equation 4**, the migration and tomography components can be expressed by combining the decomposed wavefields:

$$\mathrm{g}_{mig} = \sum\sum \mathbf{v}_u(t) * \mathbf{B}_d(t) + \mathbf{v}_d(t) * \mathbf{B}_u(t) \tag{6}$$

$$\mathrm{g}_{tomo} = \sum_{s}^{}\sum_{r}^{} \mathbf{v}_u(t) * \mathbf{B}_u(t) + \mathbf{v}_d(t) * \mathbf{B}_d(t) \tag{7}$$

However, the explicit decomposition method in the $f$-$k$ domain requires storing the wavefield at every time step, which leads to high memory cost. Figure 2 depicts the wavefield decomposition performed in $f$-$k$ domain, where repetitive 2D FFT and IFFT computations on every $z$-$t$ wavefield slice are required. In particular, this limitation becomes critical in shallow seismic applications, where high-frequency content requires dense temporal samples. To overcome this, we adopted an implicit

wavefield decomposition method proposed by Lian *et al.*,[35] which can be implemented in the time-space domain. Implicit decomposition method constructs a complex wavefield by applying Hilbert transforms in both the time and depth directions. For a wavefield $u(z,t)$, the complex wavefield $\tilde{u}(z,t)$ calculated by applying the temporal Hilbert transform $\mathcal{H}_t$ is defined as follows:[41]

$$\tilde{u}(z,t) = u(z,t) + i\mathcal{H}_t\left[u(z,t)\right] \tag{8}$$

By applying an additional Hilbert transform $\mathcal{H}_z$ in the depth direction, the wavefield can be separated by:[35]

$$\begin{aligned} u_u(z,t) &= \frac{1}{2}\left[\tilde{u}(z,t) + i\mathcal{H}_z\left[\tilde{u}(z,t)\right]\right] \\ u_d(z,t) &= \frac{1}{2}\left[\tilde{u}(z,t) - i\mathcal{H}_z\left[\tilde{u}(z,t)\right]\right] \end{aligned} \tag{9}$$

Therefore, based on **Equation 9**, **Equations 6** and **7** can be rewritten as follows:

$$\mathrm{g}_{mig} = \sum_s\sum_r 2\left[\mathbf{v}(t) * \mathbf{B}(t) + \mathcal{H}_z\left[\mathbf{v}(t)\right] * \mathcal{H}_z\left[\mathbf{B}(t)\right]\right] \tag{10}$$

$$\mathrm{g}_{tomo} = \sum_s\sum_r 2\left[\mathbf{v}(t) * \mathbf{B}(t) - \mathcal{H}_z\left[\mathbf{v}(t)\right] * \mathcal{H}_z\left[\mathbf{B}(t)\right]\right] \tag{11}$$

Based on **Equations 10** and **11**, the process of combining the up/down going wavefields (**Equations 6** and **7**) can be omitted. The migration and tomography components can be directly computed within the adjoint modeling.

### 2.3. Direct envelope inversion (DEI)

Wu and Chen[31] proposed a direct envelope inversion (DEI) based on the energy scattering theory[42] to mitigate cycle skipping caused by lack of low-frequency components. With the strong scattering condition, the objective function for DEI focuses on the mismatch between the envelope transformed modeled data ($\mathbf{Env}[u]$) and observed data ($\mathbf{Env}[d]$):

$$J_{DEI}(\mathbf{m}) = \frac{1}{2}\sum_s\sum_r\int\left[\mathbf{Env}\left[u(t,m,r;s)\right] - \mathbf{Env}\left[d(t,r;s)\right]\right]^2 dt \tag{12}$$

where envelope transformed modeled data and observed data can be expressed as follows:

$$\begin{aligned} Env\left[u(t)\right] &= \sqrt{u^2(t) + \mathcal{H}[(u(t)]^2} \\ Env\left[d(t)\right] &= \sqrt{d^2(t) + \mathcal{H}[(d(t)]^2} \end{aligned} \tag{13}$$

The gradient for DEI can be derived from **Equation 12** with respect to model parameter $\mathbf{m}$:

$$\frac{\partial J_{DEI}(\mathbf{m})}{\partial \mathbf{m}} = \sum_s\sum_r\int\frac{\partial \mathbf{Env}\left[u(t,r;s)\right]}{\partial \mathbf{m}} \cdot \mathrm{R}_{DEI}(t,r;s) \tag{14}$$

where $(\partial\mathbf{Env}\left[u(t)\right]/\partial \mathrm{m})$ is defined as the partial derivative envelope wavefield, and $\mathrm{R}_{DEI} = \boldsymbol{Env}\left[u(t)\right] - \mathbf{Env}\left[d(t)\right]$ denotes the direct envelope residual. In the strong scattering approximation, the direct envelope virtual source can be written as follow:[31]

$$v_{DEI}(x,t) = \frac{1}{c(x)}\mathrm{G}_E(x,t;x_s) \tag{15}$$

where $\mathrm{G}_E$ denotes the envelope transformed Green's operator. Therefore, the reformulated DEI gradient is given by:

$$\mathrm{g}_{DEI} = \sum_s\sum_r \mathbf{v}_{DEI}(t) * \mathbf{B}_{DEI}(t) \tag{16}$$

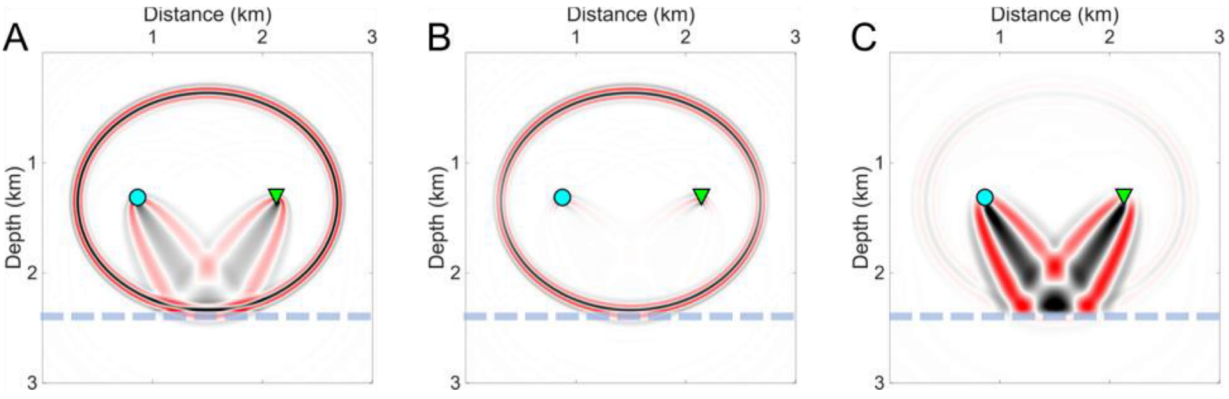



**Figure 1.** Comparison of the reflection kernels: (A) full reflection kernel, (B) migration, and (C) tomography component. The reflector, sources, and receivers are indicated by the dashed lines, cyan circles, and green triangles, respectively.

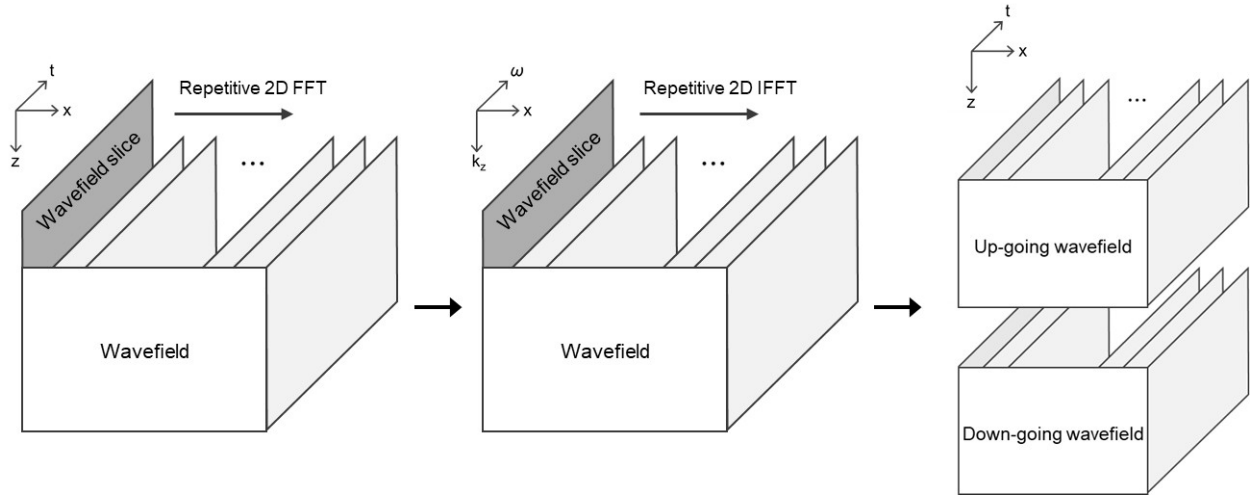



**Figure 2.** Schematic diagram of wavefield decomposition in $f$-$k$ domain.
Abbreviations: 2D FFT: Two-dimensional fast Fourier transform; 2D IFFT: Two-dimensional inversefast Fourier transform.

where $\mathbf{v}_{DEI}(t)$ is the virtual source wavefield for DEI and $\mathbf{B}_{DEI}(t)$ is the adjoint source wavefield computed by propagating the time-reversed direct envelope residual.

### 2.4. Direct envelope–based RFWI (DE-RFWI)

Reflection full waveform inversion (RFWI) can reconstruct long-wavelength velocity models from reflection data by selectively utilizing tomography components. However, conventional RFWI suffers from cycle skipping when the data have insufficient low-frequency content.[25,28] To overcome this limitation, we propose the direct envelope-based RFWI (DE-RFWI) that incorporates the direct envelope into the RFWI. In coastal seismic surveys, operating sources that generate low-frequency energy is technically challenging. Under such shallow marine conditions, incorporating the direct envelope, which emphasizes energy distribution rather than waveform phase, allows long-wavelength information to be stably estimated from reflection data lacking low-frequency components.

Figure 3 shows the two-stage workflow of DE-RFWI process. The workflow consists of the first stage that creates reflectors and the second stage that updates the long-wavelength velocity models. By using the Hilbert transform based decomposition method, the desired component at each stage can be directly obtained during the adjoint modeling without additional 2D FFT operations. Since subsurface information is usually limited, the initial velocity model does not contain reflectors, which hinders the formation of the rabbit-ear-shaped tomography kernel. Therefore, in the first stage, the migration component in **Equation 10** is used to update a temporary short-wavelength model. The short-wavelength model is used as the initial model for the second stage.

In the second stage, an envelope-based objective function is employed to recover the long-wavelength velocity models, and the tomography components are extracted accordingly. According to the **Equations 11** and **16**, the tomography component of DE-RFWI can be expressed as follows:

$$g_{tomo} = \sum_{s}\sum_{r} 2\left[\mathbf{v}_{DEI}(t) * \mathbf{B}_{DEI}(t) - \mathcal{H}_z\left[\mathbf{v}_{DEI}(t)\right] * \mathcal{H}_z\left[\mathbf{B}_{DEI}(t)\right]\right] \quad (17)$$

The tomography component is accumulated into the initial model at each iteration, whereas the short-wavelength temporary model is discarded. The updated long-wavelength model is used as the initial model for the first stage of the next iteration.

## 3. Numerical examples

### 3.1. Synthetic test

In this part of numerical example, we conducted a synthetic test on the survey conditions of a short-offset marine acquisition. This test was carried out to verify the performance of the proposed DE-RFWI and to compare the results with obtained from conventional FWI and RFWI. The velocity model used in this test has a 30 m water depth and includes an ellipse-shaped anomaly (Figure 4A). The receiver offset range is 18–87 m, with a group interval of 3 m. Assuming a vessel speed of 4 knots and a shooting interval of 2 s, the shot interval is 4 m, and the recording time is 0.3 s. These survey parameters were considered to simulate the shallow marine survey conditions. The initial velocity model is shown in Figure 4B, and a Ricker wavelet with a cutoff frequency of 100 Hz was used as the source

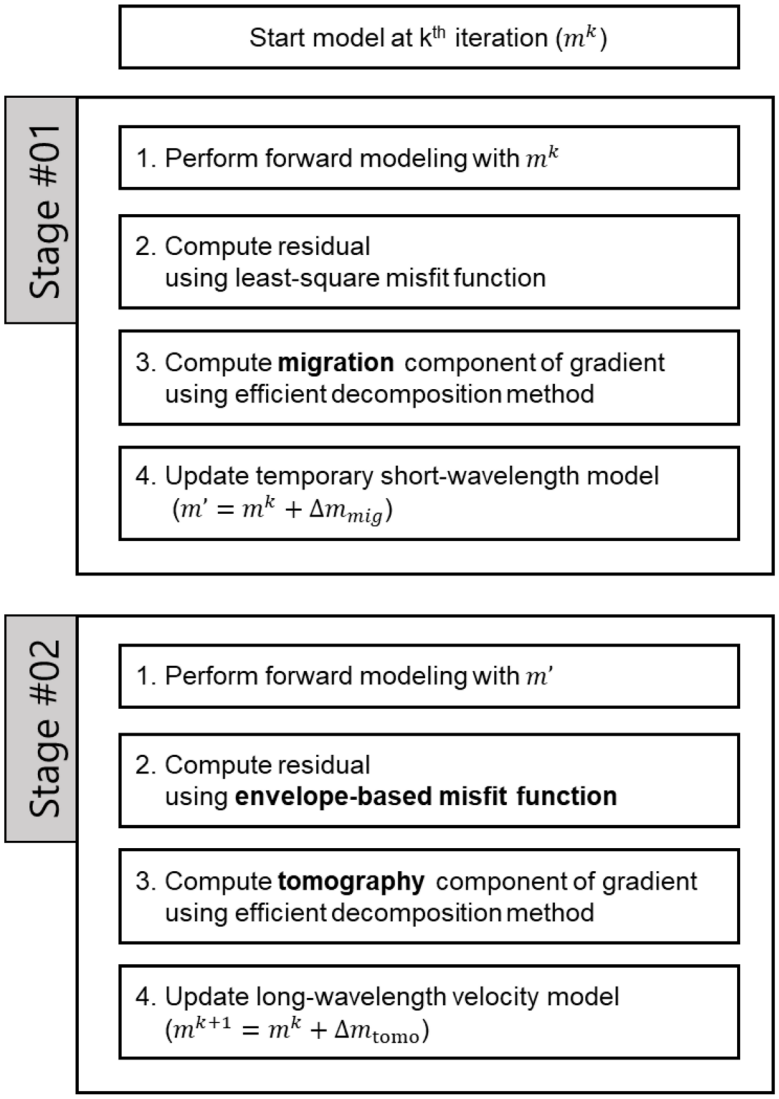



**Figure 3.** Two-stage workflow of the DE-RFWI process
Abbreviation: DE-RFWI: Direct envelope–based reflection full waveform inversion.

wavelet (Figure 5). The bandwidth and recording time in this test are consistent with typical shallow high-resolution survey parameters and the geometry of the field data. Forward modeling was carried out using the 2D acoustic wave equation under the assumption of constant density.

Since the maximum offset range is insufficient to record wide-aperture signals, most of the recorded energy consists of reflections. To verify the DE-RFWI, we compared the performance with that of conventional FWI and RFWI. Figure 6 depicts the comparison of the inversion results. For FWI, the lack of wide-aperture information leads to updates that are primarily confined to reflector boundaries, and the reflector velocities are not completely recovered (Figure 6A). RFWI also fails to reconstruct the long-wavelength velocity (Figure 6B), because the limited offset causes the migration component to dominate, leaving the separated tomography component would be still dominant by migration energy. However, the DE-RFWI yields a stable long-wavelength velocity (Figure 6C). Especially, whereas RFWI produces localized updates near the upper and lower edges of the anomaly, DE-RFWI achieves a more spatially distributed and consistent update throughout the model.

These results are also demonstrated on the velocity profiles (Figure 7). DE-RFWI provides a more stable estimation of the long-wavelength velocity compared with

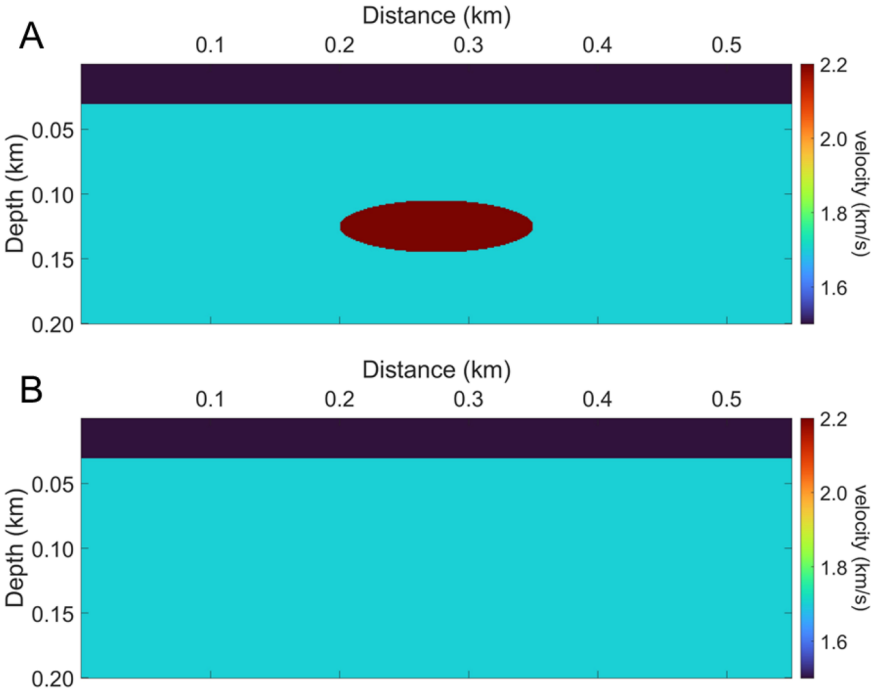



**Figure 4.** (A) True velocity model and (B) initial velocity model for synthetic test

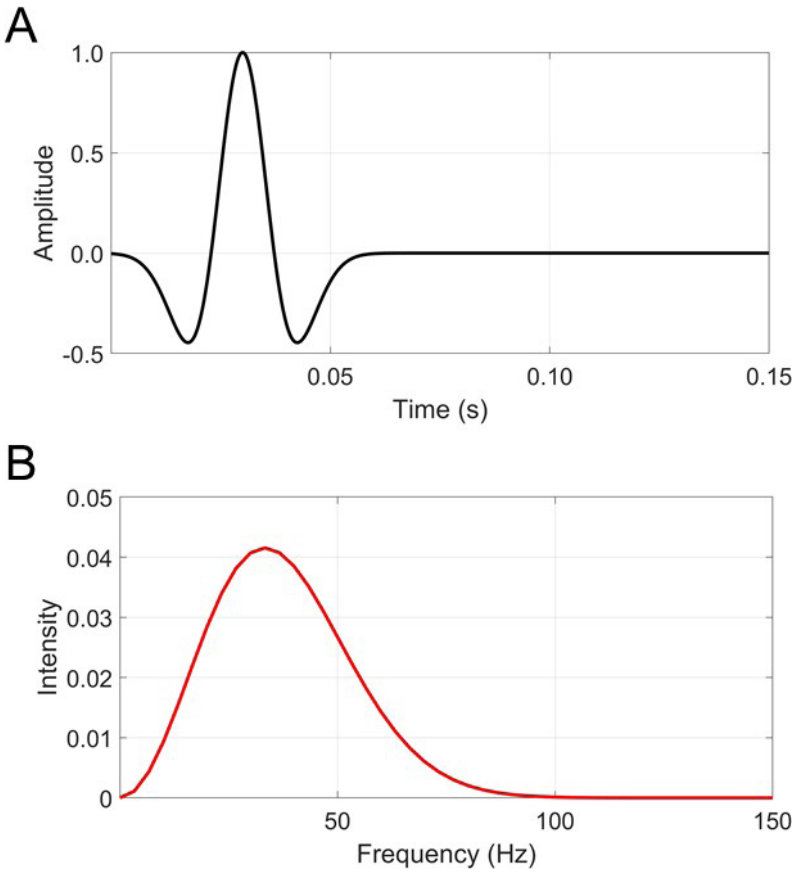



**Figure 5.** Ricker wavelet with a cutoff frequency of 100 Hz used in (A) synthetic test, and (B) its frequency spectrum

the FWI and RFWI results. Because FWI and RFWI fail to recover the background velocity, the updates occur at position shallower than the actual lower reflector, which is a consequence of the waveform-matching objective functions. However, DE-RFWI was able to reconstruct the long-wavelength velocity. Therefore, the synthetic example on the shallow marine survey demonstrates that the proposed DE-RFWI method can yield a more reliable long-wavelength velocity model with limited offsets.

### 3.2. Field data application

To verify the field applicability, we conducted DE-RFWI to seismic data acquired in Yeongil Bay, South Korea. Figure 8 shows the survey area, where water depth is less than 30 m. In the coastal region, the shallow towing depth of the source and receivers makes the direct signal highly susceptible to contamination by irregular sea-surface reflections. Thus, source estimation based on direct signals becomes unstable. In this study, we estimated the source signature using seafloor reflections, which can be recorded more reliably at short offsets. However, the use of seafloor reflections requires a reliable near-seafloor velocity model. Accordingly, we adopted the velocity model obtained from refraction tomography conducted in the same survey area.

The source equipment was an air-gun manufactured by Bolt with a chamber volume of 10 $in^3$. It was fired in a time-

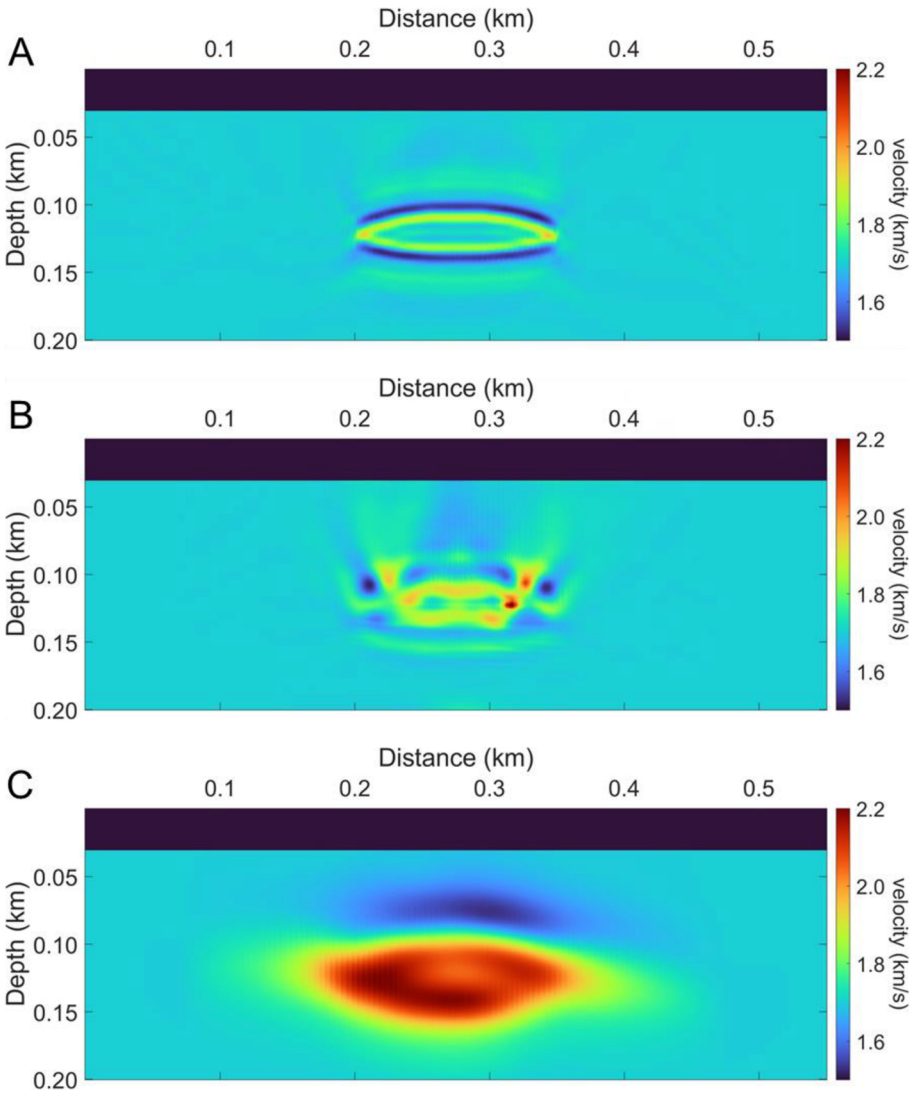



**Figure 6.** Comparison of results on (A) FWI, (B) RFWI, and (C) DE-RFWI
Abbreviations: DE-RFWI: Direct envelope–based reflection full waveform inversion; FWI: Full waveform inversion; RFWI: Reflection full waveform inversion.

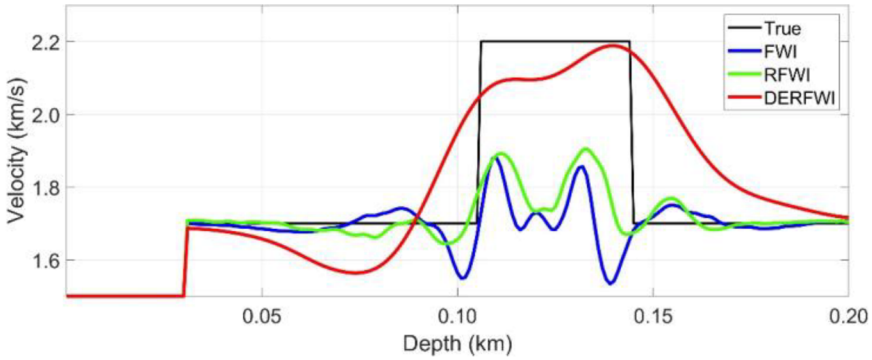



**Figure 7.** Comparison of velocity profiles extracted at the distance of 0.27 km: the true velocity (black), FWI (blue), RFWI (green), and DE-RFWI (red) Abbreviations: DE-RFWI: Direct envelope–based reflection full waveform inversion; FWI: Full waveform inversion; RFWI: Reflection full waveform inversion.

shooting method at 2 s intervals, corresponding to a shot spacing of approximately 4 m at an average vessel speed of about 4 knots. Totally, 2335 shots were recorded, of which 1000 shots were used for inversion. The receiver was a 24-channel streamer manufactured by SIG, with a group interval of 3.125 m. The offset range was about 18–90 m. The recording length was 0.2 s, and the sampling interval was 100 μs.

Pre-processing is essential to ensure the signal quality before applying field data to inversion. The trigger time between the source controller and receiver unit directly affects the temporal alignment of the recorded waveforms. A mismatch of trigger time can cause events to be delayed or recorded prematurely, leading to incorrect interpretation. In air-gun surveys, a gun-delay exists between the trigger signal generated by the air-gun controller and the actual firing, and the delay was compensated to synchronize the trigger time.

The field data contain various mechanical and electrical noises, requiring additional noise suppression procedures. In this study, high frequency noise components above 200 Hz were removed through low-pass filtering. The direct signals were muted because direct waves can be easily contaminated by irregular sea-surface reflections in shallow coastal zones. Figure 9 presents the near-offset gather and representative common shot gather after the pre-processing, and complex reflection events can be clearly observed.

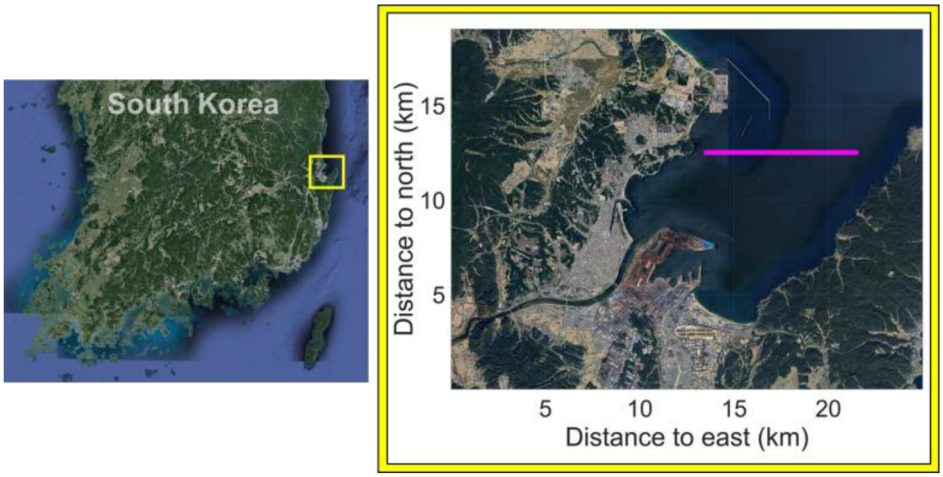



**Figure 8.** Survey area in Yeongil Bay, Sourth Korea. The survey line is indicated by a magenta line.

For field data inversion where subsurface information is limited, an initial velocity model that linearly increases with depth is widely used. In this study, we modified model to enhance the stability of the inversion by incorporating the observed characteristics of data into the initial model. Owing to the short offset and high frequency bandwidth, the data are of high resolution (Figure 9). The near-offset gather data reveals a sedimentary structure beneath the seafloor after distance of 1.8 km (Figure 9A). The two-way travel time of the reflection occurring below the seafloor reflection is approximately 0.04 s. The shallow subsurface in this region is known to consist of marine sand deposits.[43] Accordingly, instead of a simple linearly increasing velocity model, we constructed the initial model by incorporating the geomorphological features observed in the data (Figure 10A).

We carried out inversion using the modified linearly increasing model (Figure 10A) and the tomography model (Figure 10B) as initial model. Figure 11 shows the inversion results for each initial model. Although the tomography model that contains travel-time information is a relatively favorable velocity model, both results exhibit similar trends. This demonstrates that the proposed DE-RFWI can stably reconstruct long-wavelength velocity even under poor initial models.

According to the inversion results, the vertical and horizontal velocity variations are improved (Figure 11). In particular, the complex velocity structure shallower than 0.12 km is successfully recovered. Moreover, the distribution of low-velocity zones around the distance of 3.0 km and 3.5 km indicates enhanced structural resolution relative to the initial model (Figure 11B). The results of inversion demonstrate that the DE-RFWI can effectively and stably update long-wavelength velocity models when applied to short-offset data acquired in shallow water depth. However, the velocity structure below 0.1 km depth may be inaccurate. It is because the maximum offset range of the field data (0.1 km) is insufficient to adequately constrain deeper velocities, and the relatively small chamber size of air-gun may have caused signal attenuation.

To validate the result of DE-RFWI, a comparison of the observed and modeled data generated from Figure 11B was conducted using representative traces. In shallow marine conditions, the recorded data are dominated by strong reflections (Figure 9). However, since the DE-RFWI reconstructs the long-wavelength components of the velocity model, direct amplitude-based comparison is not appropriate. The long-wavelength components mainly affect travel time; therefore, we evaluated the inversion result by analyzing the arrival-time alignment between the observed and modeled data.

Figure 12 presents the windowed and their enveloped traces to analyze the arrival time alignment. For focus on the later signals, a time window after the seafloor primary reflection was applied (0.065–0.16 s). Figure 12A shows the comparison of observed and modeled traces at a shot position of 1.82 km, presented in relative amplitude. This location corresponds to the interval where sedimentary structures begin to appear from the survey line of 1.9 km (Figure 9). The relative amplitude of the signals observed at approximately 0.08 s is reasonably reproduced in the modeled trace. Figure 12B presents the envelope traces of Figure 12A. The blue boxes indicate the envelope peaks of the observed data. Comparison with the modeled envelope

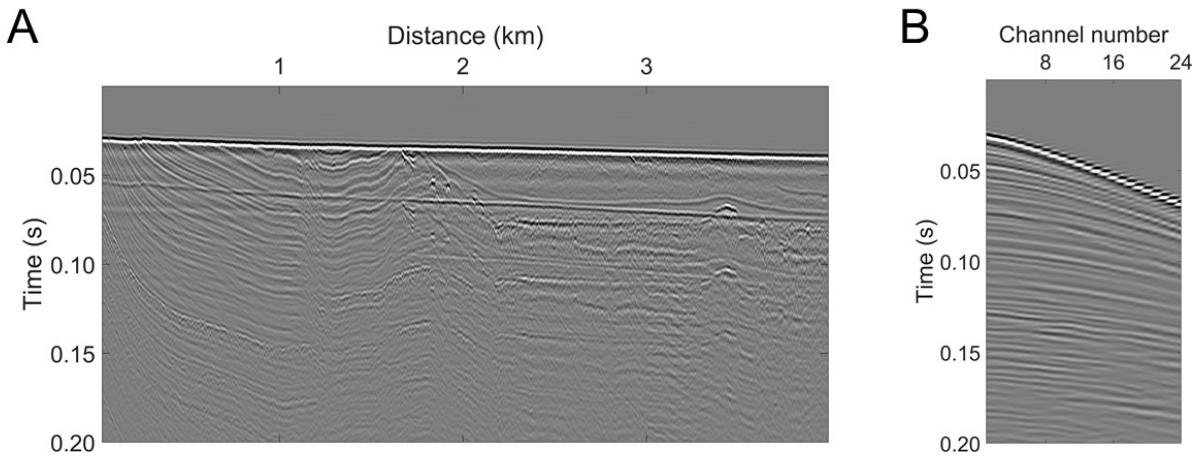



**Figure 9.** (A) Pre-processed near-offset gather and (B) representative common shot gather of field seismic data

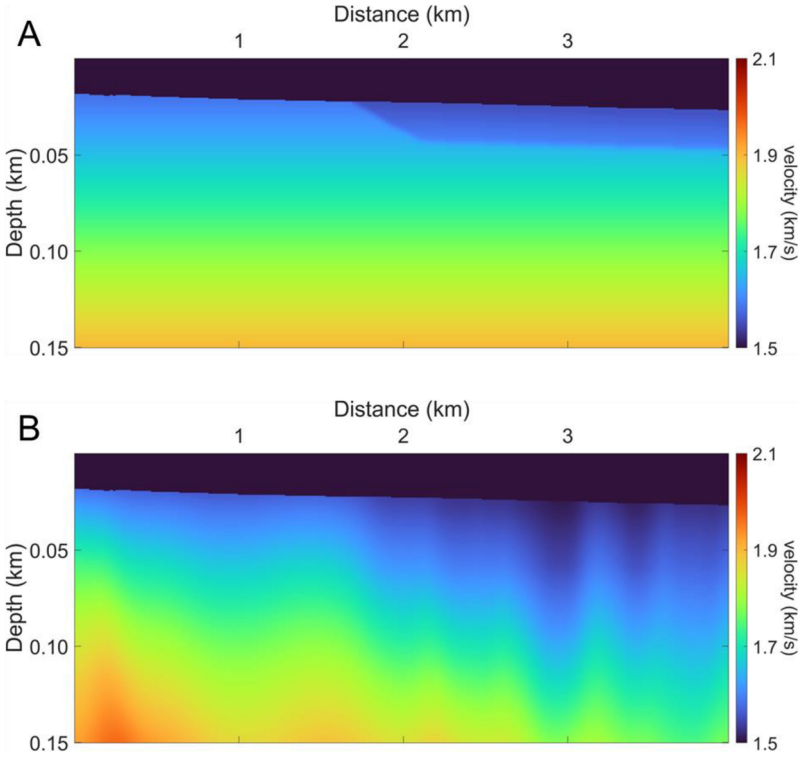



**Figure 10.** Comparison of the initial velocity models: (A) modified linearly increasing model, and (B) tomography model

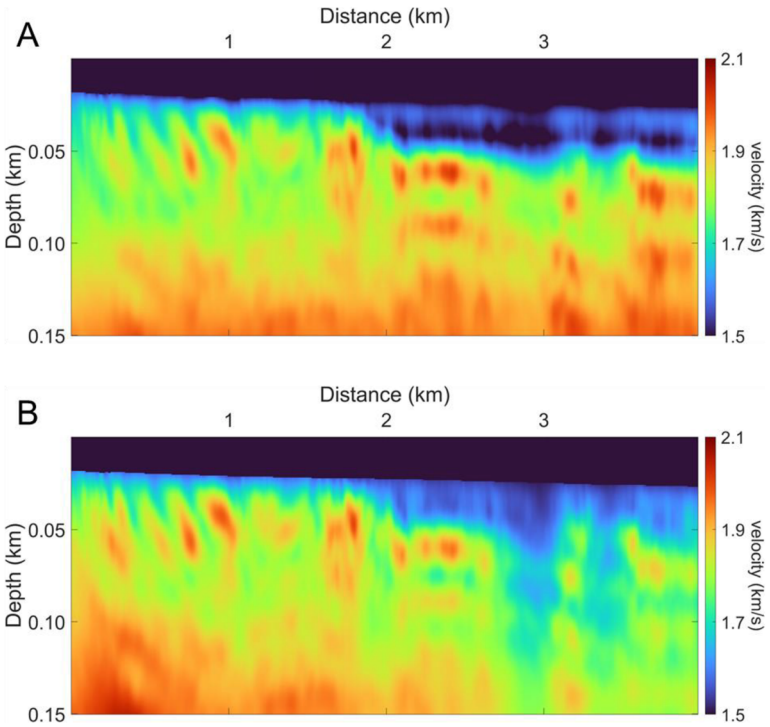



**Figure 11.** Comparison of the inverted velocities using initial model as (A) modified linearly increasing model (Figure 9A) and (B) tomography model (Figure 9B)

peak demonstrates that the arrival time of the main peaks is generally aligned.

Figure 12C displays the trace comparison of a shot position of 2.9 km. This location lies in the area influenced by the sedimentary structure. Compared with Figure 12A, the reference reflection event is observed at delayed time by the structural effect shown in inverted result (Figure 11). The reproduced relative amplitude trend can be observed (Figure 12C). Figure 12D depicts the envelope transformed data of Figure 12C. The envelope peaks within the 0.09–0.12 s are generally aligned. However, discrepancies can be detected after 0.12 s.

In this section, the result of DE-RFWI were evaluated based on relative amplitudes of windowed traces and arrival time alignment in the envelope domain. The relative amplitudes and peak-to-peak alignment are generally consistent, although additional peaks that are not clearly observed in the field data are presented in the modeled result.

Since the well-log data is not available for the Yeongil Bay field data, the inverted velocity should be verified by indirect approach. To verify the inverted velocity, a comparison of reverse time migration (RTM) was performed. Figure 13A demonstrates the RTM result using the initial velocity model in Figure 10B, whereas Figure 13B presents the RTM result derived from the inverted velocity model in Figure 11B. The green arrows indicate the improvements in the migration results. The RTM result using inverted velocity with DE-RFWI exhibits a markedly improved continuity of reflection events. Especially, the horizontal reflection around a distance of 2 km is imaged in a flattened and coherent event. These enhanced imaging results demonstrate that the proposed DE-RFWI approach can reliably and accurately reconstruct long-wavelength velocity models when applied to field data.

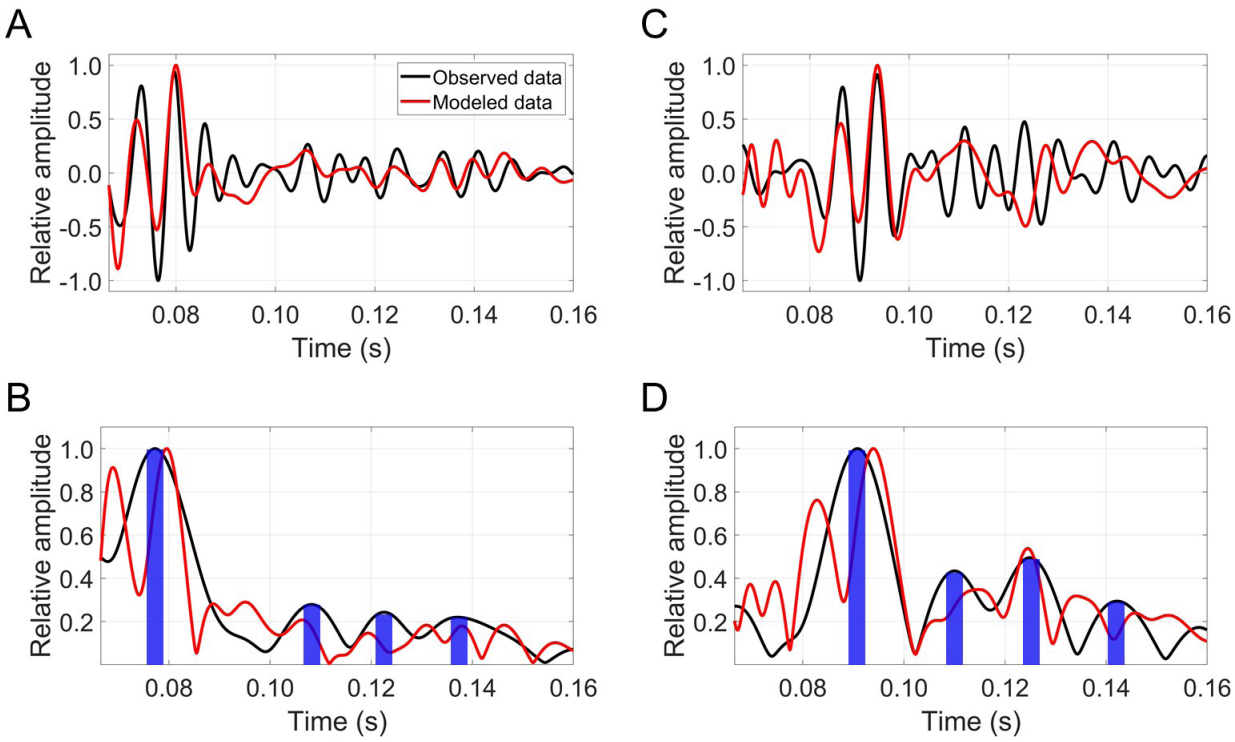



**Figure 12.** Trace comparison for relative amplitude trend and arrival-time alignment analysis: (A) First channel traces of shot located at 1.82 km, (B) envelope traces of (A), (C) first channel traces of shot located at 2.9 km, and (D) envelope traces of (C). Blue vertical bars indicate the envelope peak positions of the observed data.

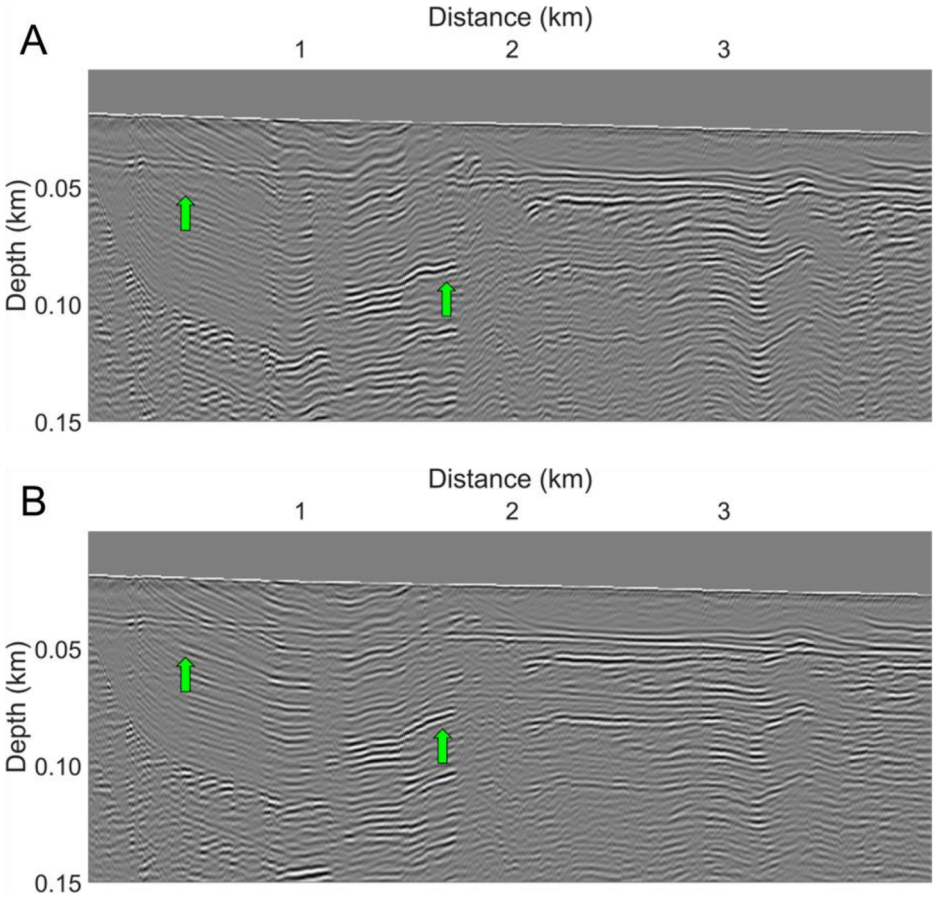



**Figure 13.** Comparison of the migration results. (A) Using the initial tomography model (Figure 10B) and (B) the inverted velocity model using DE-RFWI (Figure 11B). The green arrows indicate representative improvements.
Abbreviations: DE-RFWI: Direct envelope–based reflection full waveform inversion.

## 4. Conclusion

Shallow marine seismic data is characterized by short offsets and reflection dominance. Although RFWI has been proposed to recover long-wavelength velocity from reflection data, it remains susceptible to cycle skipping when low-frequency components are insufficient. To overcome this, we propose DE-RFWI, which incorporates the direct envelope into the RFWI to contribute to the reconstruction of long-wavelength velocity. Wavefield decomposition method to separate gradient into migration and tomography component performed in frequency-wavenumber domain requires repeated 2D FFT and storing of source and receiver wavefield at every time step, leading to high computational cost. This limitation is particularly critical in shallow marine seismic data. Therefore, we employed a Hilbert transform-based implicit decomposition method performed in time-space domain. The results of numerical examples to both synthetic and field data from Yeongil Bay, South Korea demonstrate that DE-RFWI successfully reconstructs long-wavelength velocity models, even with poor initial models. In the field data example, the inversion results show consistent arrival-time alignment and related amplitude trends in later time. The velocity model obtained from DE-RFWI also improves reflector continuity and resolution in RTM images. These results indicate that DE-RFWI is an effective approach for building reliable long-wavelength velocity models in shallow marine seismic data. In addition, proposed method is expected to support high-resolution velocity model and contribute to geotechnical assessment and engineering design in marine environments.

## Acknowledgments

We thank Dr. Sumin Kim and Principal Engineer Gabseok Seo for their insightful comments and technical expertise, which greatly improved the quality of this article.

## Funding

This work was supported by the Basic Research Project "Development of operation management infrastructure for TAMHAE3 and seamless seismic technology connecting coastal areas (26-3322)" of the Korea Institute of Geoscience and Mineral Resources (KIGAM) funded by the Ministry of Science and ICT of Korea.

## Conflict of interest

The authors declare that there are no conflicts of interest or competing interests.

## Author contributions

*Conceptualization:* Kyoungmin Lim, Jiho Ha
*Formal analysis:* Wookeen Chung, Jungkyun Shin
*Investigation:* Kyoungmin Lim, Jiho Ha, Jungkyun Shin
*Methodology:* Kyoungmin Lim, Jiho Ha
*Supervision:* Jiho Ha
*Validation:* Wookeen Chung
*Visualization:* Kyoungmin Lim
*Writing–original draft:* Kyoungmin Lim, Jiho Ha
*Writing–review & editing:* All authors

## Availability of data

All data analyzed have been presented in the paper.

## References

1. Prager BT, Caughey DA, Poeckert RH. Bottom classification: operational results from QTC VIEW. In: *Proceedings of the Challenges of Our Changing Global Environment.* OCEANS '95 MTS/IEEE; October 9-12, 1995; San Diego, CA, USA. 1995;3:1827-1835.

   doi: 10.1109/OCEANS.1995.528859

2. Mair E, Parnum I, Schut T. Using multibeam echo-sounder backscatter to characterise the seafloor of Davis Harbour, Antarctica: Preliminary results. In: *Proceedings of the Acoustics 2012 – Fremantle: Acoustics, Development, and the Environment.* The 2012 Conference of the Australian Acoustical Society; November 21-23, 2012; Fremantle, Australia. Australian Acoustical Society; 2012.

3. Viala C, Lamouret M, Abadie A. Seafloor classification using a multibeam echo sounder: A new rugosity index coupled with a pixel-based process to map Mediterranean marine habitats. *Appl Acoust*. 2021;179:108067.

   doi: 10.1016/j.apacoust.2021.108067

4. Brookshire BN Jr, Landers FP, Stein JA. Applicability of ultra-high-resolution 3D seismic data for geohazard identification at mid-slope depths in the Gulf of Mexico: Initial results. *Underw Technol.* 2015;32(4):271-278.

   doi: 10.3723/ut.32.271

5. Shin J, Ha J, Kang NK, Kim HD, Kim CS. Development of a portable 3D seismic survey system for nearshore surveys and the first case study offshore Pohang, South Korea. *Mar Geophys Res*. 2021;42(4):34.

   doi: 10.1007/s11001-021-09453-x

6. Luo Y, Schuster GT. Wave-equation traveltime inversion. *Geophysics*. 1991;56(5):645-653.

   doi: 10.1190/1.1443081

7. Zelt CA, Smith RB. Seismic traveltime inversion for 2-D crustal velocity structure. *Geophys J Int*. 1992;108(1):16-34.

   doi: 10.1111/j.1365-246x.1992.tb00836.x

8. Ravaut C, Operto S, Improta L, Virieux J, Herrero A, Dell'Aversana P. Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt. *Geophys J Int*. 2004;159(3):1032-1056.

   doi: 10.1111/j.1365-246x.2004.02442.x

9. Pichot T, Delescluse M, Chamot-Rooke N, *et al*. Deep crustal structure of the conjugate margins of the SW South China Sea from wide-angle refraction seismic data. *Mar Pet Geol*. 2014;58:627-643.

   doi: 10.1016/j.marpetgeo.2013.10.008

10. Mora P. Inversion = migration + tomography. *Geophysics*. 1989;54(12):1575-1586.

    doi: 10.1190/1.1442625

11. Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics. *Geophysics*. 2009;74(6):WCC1-WCC26.

    doi: 10.1190/1.3238367

12. Zhou W, Brossier R, Operto S, Virieux J. Full waveform inversion of diving and reflected waves for velocity model building with impedance inversion based on scale separation. *Geophys J Int*. 2015;202(3):1535-1554.

    doi: 10.1093/gji/ggv228

13. Xu S, Wang D, Chen F, Lambaré G, Zhang Y. Inversion on Reflected Seismic Wave. In: *Proceedings of the SEG Technical Program Expanded Abstracts 2012.* 82th SEG Annual International Meeting; November 4-9, 2012; Las Vegas, Nevada, USA. Society of Exploration Geophysicists; 2012:1-7.

    doi: 10.1190/segam2012-1473.1

14. Tang Y, Lee S, Baumstein A, Hinkley D. Tomographically enhanced full wavefield inversion. In: *Proceedings of the SEG Technical Program Expanded Abstracts 2013.* 83th SEG Annual International Meeting; September 22-27, 2013; Houston, Texas, USA. Society of Exploration Geophysicists; 2013:1037-1041.

    doi: 10.1190/segam2013-1145.1

15. Alkhalifah T. Scattering-angle based filtering of the waveform inversion gradients. *Geophys J Int*. 2015;200(1):363-373.

doi: 10.1093/gji/ggu379

16. Wang F, Donno D, Chauris H, Calandra H, Audebert F. Waveform inversion based on wavefield decomposition. *Geophysics*. 2016;81(6):R457-R470.

doi: 10.1190/geo2015-0340.1

17. Yao G, da Silva NV, Warner M, Kalinicheva T. Separation of Migration and Tomography Modes of Full-Waveform Inversion in the Plane Wave Domain. *J Geophys Res Solid Earth*. 2018;123(2):1486-1501.

doi: 10.1002/2017jb015207

18. Guo Q, Alkhalifah T. Elastic reflection-based waveform inversion with a nonlinear approach. *Geophysics*. 2017;82(6):R309-R321.

doi: 10.1190/geo2016-0407.1

19. Li Y, Alkhalifah T. Multi-parameter reflection waveform inversion for acoustic transversely isotropic media with a vertical symmetry axis. *Geophys Prospect*. 2020;68(6):1878-1892.

doi: 10.1111/1365-2478.12966

20. Wu S, Wang T, Cheng J. Second-order optimization for multiparameter reflection waveform inversion in acoustic VTI media. *Geophys J Int*. 2024;236(1):249-269.

doi: 10.1093/gji/ggad406

21. Song C, Alkhalifah T. A reflection-based efficient wavefield inversion. *Geophysics*. 2021;86(4):R497-R508.

doi: 10.1190/geo2019-0664.1

22. Wang T, Cheng J, Geng J. Reflection full waveform inversion with second-order optimization using the adjoint-state method. *J Geophys Res Solid Earth*. 2021;126(8):e2021.

doi: 10.1029/2021jb022135

23. Xu W, Hu G, He B, Du Z. Seismic reflection waveform inversion based on Gauss–Newton optimization. *J Geophys Eng*. 2022;19(4):846-862.

doi: 10.1093/jge/gxac053

24. Xu S, Chen F, Lambaré G, Zhang Y. Full waveform inversion of reflected seismic data. *J Seism Explor*. 2013;22(5):449-462.

25. Wang S, Chen F, Zhang H, Shen Y. Reflection-based full waveform inversion (RFWI) in the frequency domain. In: *Proceedings of the SEG Technical Program Expanded Abstracts 2013*. 83th SEG Annual International Meeting; September 22-27, 2013; Houston, Texas, USA. Society of Exploration Geophysicists; 2013:877-881.

doi: 10.1190/segam2013-0671.1

26. Wang P, Zhang Z, Wei Z, Huang R. A demigration-based reflection full-waveform inversion workflow. In: *Proceedings of the SEG Technical Program Expanded Abstracts 2018*. 88th SEG Annual International Meeting; October 14-19, 2018; Anaheim, California, USA. Society of Exploration Geophysicists; 2018:1138-1142.

doi: 10.1190/segam2018-2997404.1

27. Yao G, Wu D, Wang SX. A review on reflection-waveform inversion. *Pet Sci*. 2020;17(2):334-351.

doi: 10.1007/s12182-020-00431-3

28. Chi B, Dong L, Liu Y. Correlation-based reflection full-waveform inversion. *Geophysics*. 2015;80(4):R189-R202.

doi: 10.1190/geo2014-0345.1

29. Chen G, Wu R, Chen S. Reflection multi-scale envelope inversion. *Geophys Prospect*. 2018;66(7):1258-1271.

doi: 10.1111/1365-2478.12624

30. Wang Y, Chi B, Dong L. Envelope normalized reflection waveform inversion. *Geophys Prospect*. 2024;73(3):895-909.

doi: 10.1111/1365-2478.13598

31. Wu RS, Chen GX. Multi-scale seismic envelope inversion using a direct envelope Frechet derivative for strong-nonlinear full waveform inversion. *arXiv*. Preprint posted online August 15, 2018. Accessed March 19, 2026.

doi: 10.48550/arXiv.1808.05275

32. Chen GX, Wu RS, Chen SC. Multiscale Direct Envelope Inversion: Algorithm and Methodology for Application to the Salt Structure Inversion. *Earth Space Sci*. 2019;6(1):174-190.

doi: 10.1029/2018EA000453

33. Luo J, Wu RS, Chen G. Angle domain direct envelope inversion method for strong scattering velocity and density estimation. *IEEE Geosci Remote Sens Lett*. 2020;17(9):1508-1512.

doi: 10.1109/LGRS.2019.2950471

34. Lee D, Kang SG, Kim S, Kim YS, Chung W. Efficient direct envelope inversion with excitation amplitude for strong velocity contrast model. *IEEE Trans Geosci Remote Sens*. 2024;62:1-12.

doi: 10.1109/TGRS.2024.3422978

35. Lian S, Yuan S, Wang G, Liu T, Liu Y, Wang S. Enhancing low-wavenumber components of full-waveform inversion using an improved wavefield decomposition method in the time-space domain. *J Appl Geophys*. 2018;157:10-22.

doi: 10.1016/j.jappgeo.2018.06.013

36. Jun H. Frequency-domain reflection-based full waveform inversion for short-offset seismic data. *J Appl Geophys*. 2019;164:106-116.

doi: 10.1016/j.jappgeo.2019.03.010

37. Wu Z, Alkhalifah T. Simultaneous inversion of the background velocity and the perturbation in full-waveform

inversion. *Geophysics*. 2015;80(6):R317-R329.

doi: 10.1190/geo2014-0365.1

38. Tarantola A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*. 1984;49(8):1259-1266.

doi: 10.1190/1.1441754

39. Pratt RG, Shin C, Hicks GJ. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophys J Int*. 1998;133(2):341-362.

doi: 10.1046/j.1365-246x.1998.00498.x

40. Hu LZ, McMechan GA. Wave-field transformations of vertical seismic profiles. *Geophysics*. 1987;52(3):307-321.

doi: 10.1190/1.1442305

41. Fei TW, Luo Y, Yang J, Liu H, Qin F. Removing false images in reverse time migration: The concept of de-primary. *Geophysics*. 2015;80(6):S237-S244.

doi: 10.1190/geo2015-0289.1

42. Wu RS. Multiple scattering and energy transfer of seismic waves — separation of scattering effect from intrinsic attenuation — I. Theoretical modelling. *Geophys J Int*. 1985;82(1):57-80.

doi: 10.1111/j.1365-246x.1985.tb05128.x

43. Ha J, Shin J, Lim K, Um IK, Yi B. 3D UHR seismic and back-scattering analysis for seabed and ultra-shallow subsurface classification. *Acta Geophys*. 2024;73(2):1363-1376.

doi: 10.1007/s11600-024-01423-2

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Enhanced amplitude variation with incident angle inversion with Cauchy regularization using ensemble smoother-based data assimilation

**Rui Liao and Ronghuo Dai***

Department of Statistics, School of Mathematical Sciences, China West Normal University, Nanchong, Sichuan, China

## Abstract

Amplitude variation with incident angle (AVA) inversion for partial angle stack seismic data is an extension of acoustic impedance inversion. It adopts the P-wave reflection coefficient to link seismic amplitude information with elastic parameters and incident angle. This method generalizes conventional acoustic impedance inversion to pre-stack seismic data, enabling the effective inversion of multiple elastic parameters by utilizing seismic data acquired at different incident angles, such as angle gather data or partial angle stack data. This study addresses the limitations of commonly used AVA inversion with Cauchy regularization. Due to the non-linearity of Cauchy regularization, the iterative re-weighted least squares algorithm is extensively employed. It has been shown that the accuracy of inversion results of this non-linear inversion algorithm is highly dependent on the initial solution. When geological conditions are complex, the initial solution must be close to the optimal solution to ensure accurate inversion results. To address this challenge, this study proposes combining the ensemble smoother with multiple data assimilation (ES-MDA) and AVA inversion with Cauchy sparse regularization. Specifically, the posterior mean of ES-MDA is used as the initial solution for AVA inversion with Cauchy sparse regularization. ES-MDA is a stochastic method that solves inverse problems by iteratively updating an ensemble of model realizations, yielding a solution that closely approximates the optimal solution. Practical application indicates that, compared with AVA inversion with Cauchy sparse regularization using a standard initial solution, the proposed method achieves improved accuracy in estimated elastic parameters. The research findings offer new technical approaches for seismic prediction with AVA inversion in complex reservoirs.

***Keywords:*** Seismic inversion; Ensemble smoother with multiple data assimilation; Elastic parameters; Amplitude variation with incident angle inversion; Cauchy regularization

***Corresponding author:**
Ronghuo Dai
(daironghuo@yeah.net)

**Citation:** Liao R, Dai R. Enhanced amplitude variation with incident angle inversion with Cauchy regularization using ensemble smoother-based data assimilation. *J Seismic Explor*. 2026;35(2):025500125. doi: 10.36922/JSE025500125

**Received:** December 10, 2025

**Revised:** January 10, 2026

**Accepted:** February 4, 2026

**Published online:** March 31, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Due to the advancements in global oil and gas exploration and development, exploration targets have gradually shifted from simple structural reservoirs to complex lithologic reservoirs, unconventional reservoirs, and deep-to-ultra-deep domains. Such reservoirs often exhibit pronounced heterogeneity and complex pore fluid properties, posing

significant challenges for predicting reservoir properties and identifying fluid types using traditional seismic amplitude data.

Elastic parameters that are sensitive to lithology and fluids, such as P-impedance, P-velocity, S-impedance, S-velocity, Lamé parameters, and density, can be directly estimated from seismic traces at different incidence angles using pre-stack inversion techniques. These techniques are fundamentally based on the Zoeppritz equation[1,2] and its approximations and have become a core method for accurately characterizing complex reservoirs. Among these, amplitude variation with incident angle (AVA) inversion serves as a bridge, tightly linking pre-stack data with well-log data. Its clear physical significance and relatively straightforward computation have led to widespread adoption in the industry. By combining AVA inversion from multiple angles to invert subsurface elastic parameters, the multi-solution nature of the inversion problem can be effectively reduced, enhancing the accuracy of reservoir characterization. The core of AVA inversion lies in estimating the subsurface elastic parameters. This methodology is based on the Zoeppritz equation,[2] which is complex and difficult to apply directly, giving rise to many approximate methods that simplify it from different perspectives. Aki and Richards[1] proposed linear approximations suitable for weakly heterogeneous media and small incident angles. Shuey[3] reconfigured the Aki–Richards equation using P-velocity, Poisson's ratio, and density; this version is widely used in gas reservoir characterization. Russell *et al.*[4] introduced the effect of pore fluid, directly estimating the fluid indicator factor. Xiao *et al.*[5] proposed a non-linear pre-stack AVA inversion method incorporating lithology constraints. Within a Bayesian framework, their method introduces correlations between elastic parameters to establish constraints, demonstrating enhanced resolution and stability. Bao *et al.*[6] developed an AVO inversion method by replacing approximate isotropic Zoeppritz equations with the exact isotropic Zoeppritz equations, improving the accuracy of the isotropic model. Ye *et al.*[7] analyzed errors in the Zoeppritz approximate equations using log data and demonstrated the feasibility of full-angle pre-stack density inversion through AVO modeling. As seismic exploration targets become increasingly complex and large-offset data are widely adopted, the industry's demand for AVA inversion methods applicable to strong heterogeneous media and wide incident angles has grown. To address this challenge, Zhou *et al.*[8] proposed the joint multi-wave inversion using the exact Zoeppritz equation with a modified Cauchy prior distribution. By establishing a stochastic model for reservoir reflectivity variation, Li[9] derived a reservoir-characteristics-constrained AVO equation that allows pre-stack inversion to recover fluid-sensitive parameters. Lehocki *et al.*[10] effectively improved the accuracy of density inversion using high-order approximation of the Zoeppritz equations. Sun *et al.*[11] developed a high-precision fluid factor using a frequency-varying AVO inversion based on the Russell equation to extract frequency dispersion attributes.

However, the ill-posed nature of AVA inversion remains a significant challenge. To obtain stable and geologically reasonable inversion results, regularization is required.[12] Among numerous choices, Cauchy regularization is widely used in seismic inversion.[13-15] Due to its non-linearity, the iterative re-weighted least squares (IRLS) algorithm is commonly employed.[16] The accuracy of IRLS inversion is highly dependent on the initial solution, making its selection critical in seismic inversion with Cauchy regularization. Common approaches for generating initial solutions include zero-ordinary-solution,[17] well-log interpolation,[18] and velocity analysis.[19] When geological conditions are simple and non-linearity is limited, these initial solutions produce satisfactory inversion results without significantly affecting subsequent reservoir predictions or fluid identification. However, in complex geological settings, such as lithologic pinch-out, channels, steeply inclined strata, hydrocarbon reservoirs, or reservoirs with abrupt lateral variations of elastic parameters, the quality of the inversion results can be severely compromised. In these cases, the initial solution must be close to the optimal solution to ensure accuracy.

To address this limitation of AVA inversion with Cauchy regularization, we propose combining the ensemble smoother with multiple data assimilation (ES-MDA) and AVA inversion using Cauchy sparse regularization. ES-MDA, developed by Emerick and Reynolds,[20] enhances the handling of non-linear problems through ensemble statistics and multiple iterations. As an ensemble-based Bayesian inversion algorithm, ES-MDA has been successfully applied in seismic inversion, effectively addressing the bottlenecks of traditional methods in handling non-linearity and quantifying uncertainty.[21] ES-MDA is a stochastic method that solves inverse problems by iteratively updating an ensemble of model realizations, producing a solution that approximates the optimal solution. However, seismic reservoir characterization often involves high-dimensional geophysical data. Consequently, the number of high-dimensional ensemble models for complete two-dimensional (2D) profiles or three-dimensional (3D) volumes can become prohibitively large, limiting the method's practical application.[22] In this study, we implement ES-MDA in a computationally feasible manner, applying it trace-by-trace to exploit its

ability to approximate the optimal solution. Research indicates that, when combined with ES-MDA, the proposed AVA inversion method achieves improved accuracy in estimating elastic parameters.

## 2. Methods

### 2.1. Amplitude variation with incident angle inversion with Cauchy regularization

The AVA inversion is a technique used to extract the elastic properties of subsurface rocks—such as velocity and density—from seismic data. Specifically, it analyzes how the amplitude of seismic reflections changes with angle, using pre-stack angle gathers or partially angle-stack data as the primary input. These properties encompass fundamental elastic parameters—such as P-wave velocity, S-wave velocity, and density—as well as derived attributes such as Poisson's ratio, Young's modulus, Lamé parameters, fluid factors, and other relevant attributes. By estimating these parameters, a more comprehensive understanding of the physical characteristics and structure of underground rocks can be achieved, which is crucial for oil and gas exploration and development.

Currently, AVA inversion typically involves first estimating the relative variation rates of elastic parameters from pre-stack seismic data. Subsequently, the final elastic parameters are derived through their relationships with these relative variation rates of elastic parameters. In this context, the objective function for AVA inversion is formulated based on the Aki–Richards approximation of the P-wave reflection coefficient, as shown in **Equation 1**:

$$R_{pp}(\theta) \approx a(\theta)R_{vp} + b(\theta,\gamma)R_{vs} + c(\theta,\gamma)R_{\rho} \tag{1}$$

where, $a(\theta) = \sec^2\theta$, $b(\theta,\gamma) = -8\gamma^2\sin^2\theta$, $c(\theta,\gamma) = (1-4\gamma^2\sin^2\theta)$, $\theta$ represents the incident angle, $R_{pp}(\theta)$ represents the P-wave reflection coefficient, $\gamma$ is the ratio of S-velocity to P-velocity, and $R_{V_P}$, $R_{V_S}$, and $R_{\rho}$ are the relative variation rates of $V_p$, $V_s$, and $\rho$, respectively. Here, $V_p$, $V_s$, and $\rho$ represent P-velocity, S-velocity, and density, respectively.

Using the seismic convolution model together with the approximate expression given in **Equation 1**, the forward equation for AVA inversion can be written as **Equation 2**:[23]

$$\mathbf{d} = \mathbf{Am} + \mathbf{n} \tag{2}$$

where **A** represents the forward operator incorporating the wavelet convolution operator and the approximate expression, $\mathbf{m} = [R_{Vp}, R_{Vs}, R_{\rho}]^T$ denotes the model parameters for AVA inversion, **d** represents the observations (i.e., the input seismic data), and **n** denotes the noise.

Typically, the AVA inversion is performed by minimizing a least squares objective function, regularized by a combination of Cauchy regularization and a priori model constraints to enhance both resolution and stability, as shown in **Equation 3**:

$$\min f(\mathbf{m}) = \| \mathbf{Am} - \mathbf{d} \|_2^2 + \lambda \text{Cauchy}(\mathbf{m}) + \alpha \| \mathbf{m} - \mathbf{m}_{prior} \|_2^2 \tag{3}$$

where $\mathbf{m}_{prior}$ represents the a priori model, as shown in **Equation 4**:

$$\text{Cauchy}(\mathbf{m}) = \sum_i \ln[1 + \frac{m^2(i)}{\sigma^2}] \tag{4}$$

where the parameters in the objective function are defined as follows: $\sigma$ is the deviation parameter of the model $m$ under the Cauchy criterion;[9] $\lambda$ and $\alpha$ are the regularization parameters corresponding to the Cauchy and a priori model terms, respectively; and $m(i)$ denotes the $i$-th element of $m$.

The estimation of the final elastic parameters is typically a two-step process. First, the AVA inversion objective function is solved to obtain the relative variations of the elastic parameters. Second, these relative variations are integrated using established formulas to recover the absolute elastic properties.[15,24] Alternatively, other approximate expressions can be used to directly estimate elastic parameters through AVA inversion, such as Fatti's approximation,[25] Li and Zhang's approximation,[23] and related approaches.

### 2.2. Iteratively reweighted least squares algorithm

In general, the objective function in **Equation 3** is solved using the IRLS algorithm.[16] The basic steps of the IRLS method are as follows. From **Equation 4**, the derivative of the Cauchy regularization term (i.e., the second term in **Equation 3**) can be expressed as **Equation 5**:

$$\nabla \sum_{i=1}^{m} \frac{2m(i)}{m^2(i) + \sigma^2} = 2\mathbf{Q}(\mathbf{m})\mathbf{m} \tag{5}$$

where Q(m) is a diagonal matrix whose elements are $\frac{1}{\sigma^2 + m^2(i)}$.

Hence, the solution to **Equation 3** can be written as **Equation 6**:

$$\mathbf{A}^T\mathbf{Am} + \lambda\mathbf{Q}(\mathbf{m})\mathbf{m} + \alpha\mathbf{m} = \mathbf{A}^T\mathbf{d} + \alpha\mathbf{m}_{prior} \tag{6}$$

The basic idea of IRLS is as follows: Q(m) is treated as a weight matrix. To solve **Equation 3**, IRLS iteratively performs the following damped least-squares problem,

starting from a suitable initial solution, as shown in **Equation 7**:

$$\mathbf{m}^k = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{Q}(\mathbf{m}^{k-1}) + \alpha\mathbf{I})^{-1}(\mathbf{A}^T\mathbf{d} + \alpha\mathbf{m}_{prior}) \quad (7)$$

In each iteration, Q(m) is updated using the solution from the previous iteration. The solution of **Equation 7** can be obtained using commonly adopted methods, such as the conjugate gradient algorithm. The combined effect of the Cauchy regularization term and the inherent non-linearity of the forward equation makes the IRLS algorithm prone to convergence toward a local optimum during the inversion process. Consequently, the choice of the initial solution is a critical factor in AVA inversion. Despite this limitation, the iterative framework of IRLS is simple, robust, and widely used in the industry. In this study, quality control was employed to determine the regularization parameters, which required available well-log data. During quality control, the actual well logs were regarded as the answer for the inversion of near-well seismic traces, as they represent the actual geological characteristics of underground formations. The optimal regularization parameters were determined by performing quality control at well locations, including adjusting the value of the regularization parameter set, obtaining the inversion result from the near-well seismic traces for each regularization parameter set, and selecting the set that produces the best match with the corresponding well logs. The selected regularization parameters were then applied to the inversion for other seismic traces.

### 2.3. Ensemble smoother data assimilation

To construct a reliable initial solution, ES-MDA was integrated into the AVA inversion framework. ES-MDA approximates the posterior probability distribution through ensemble statistics. The core idea of ES-MDA is to iteratively and gradually incorporate observational information into an initial set of model parameters (i.e., multiple possible models). This process ensures that the updated parameter set not only better fits the observations but also reasonably reflects parameter uncertainties. The specific steps of the ES-MDA algorithm are as follows:[22]

(i) Define the ensemble size *n*, determine the number of data assimilation iterations *K*, and the inflation coefficient $\{\alpha_i, i = 1,2,...,K\}$ for each iteration.

(ii) Samples are drawn from the prior distribution of the model parameters to form the initial sample set $\{\mathbf{m}_j, j = 1,\ldots,n\}$. Let $i$=1.

(iii) Start the iteration. Simulate *n* zero-mean Gaussian random vectors $\{\mathbf{z}_j, j = 1,\ldots,n\}$, where $\mathbf{z}_j \sim N(0,\delta^2\mathbf{I})$, is the variance and $\mathbf{I}$ is the identity matrix. Then, the simulated random vectors are added to the measured seismic data as the perturbations, as shown in **Equation 8**:

$$\tilde{\mathbf{d}}_j = \mathbf{d} + \sqrt{\alpha_i}\mathbf{£}^{\frac{1}{2}}\mathbf{z}_j, j = 1,\ldots,n \quad (8)$$

where, **Ó** is the covariance matrix of the original data d, $\tilde{\mathbf{d}}_j$ is the *j*-th perturbed data vector.

(iv) Apply the forward operator A to each model parameter $\mathrm{m}_j$ in $\{\mathbf{m}_j, j = 1,\ldots,n\}$ to calculate the predictive data $\{\mathbf{d}_j, j = 1,\ldots,n\}$.

(v) Update the model parameters. For each model parameter, update $\mathbf{m}_j$ in the ensemble model $\{\mathbf{m}_j, j = 1,\ldots,n\}$,

$$\tilde{\mathbf{m}}_j = \mathbf{m}_j + \mathbf{£}_{m,d}(\mathbf{£}_{d,d} + \alpha_i\mathbf{£})^{-1}(\tilde{\mathbf{d}}_j - \mathbf{d}_j) \quad (9)$$

where, $\tilde{\mathbf{m}}_j$ is the *j*-th model after updating, $\mathbf{Ó}_{m,d}$ is the covariance matrix between the ensemble models and predictive data, and $\mathbf{Ó}_{d,d}$ is the covariance matrix of predictive data, as shown in **Equations 10** and **11**:

$$\mathbf{£}_{m,d} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{m}_j - \bar{\mathbf{m}})(\mathbf{d}_j - \bar{\mathbf{d}})^T \quad (10)$$

$$\mathbf{£}_{d,d} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{d}_j - \bar{\mathbf{d}})(\mathbf{d}_j - \bar{\mathbf{d}})^T \quad (11)$$

and, $(\Sigma_{d,d} + \alpha_i\Sigma)^{-1}$ is the so-called Kalman gain matrix. Here, $\bar{\mathbf{d}}$ and $\bar{\mathbf{m}}$ are the mean vectors of $\{\mathbf{d}_j, j = 1,\ldots,n\}$ and $\{\mathbf{m}_j, j = 1,\ldots,n\}$.

(vi) Replace the initial ensemble model $\{\mathbf{m}_j, j = 1,\ldots,n\}$ with the updated model $\{\tilde{\mathbf{m}}_j, j = 1,\ldots,n\}$.

(vii) Let $i$=$i$+1 and repeat steps (iii)–(vi) until the *K*-th iteration is completed, resulting in the final set of parameter samples $\{\mathbf{m}_j, j = 1,\ldots,n\}$ and derive the estimations of the posterior distribution of the model parameters.

In ES-MDA, the quality of results depends on the number of assimilation iterations $n_\alpha$, the inflation coefficient $\alpha_i$, and the Gaussian perturbations variance $\delta^2$. Traditional ES-MDA approaches usually determine these parameters based on experience selection. Following the quality control strategy described earlier, we applied the same approach here to obtain optimal values for these parameters in ES-MDA.

Ensemble smoother with multiple data assimilation refines each ensemble model by minimizing the mismatch between observed data and model predictions, resulting in a series of model realizations that align with the observed outcomes. This method provides a stochastic solution to

an inverse problem by sequentially updating an ensemble of model variables through simultaneous assimilation of measurement data. Its solution is a good approximation of the optimal solution. Therefore, the posterior mean of the ES-MDA results was adopted as the initial solution of AVA inversion with Cauchy regularization.

However, seismic reservoir characterization often involves geophysical data of very high dimensionality. Consequently, the number of high-dimensional ensemble models for the whole 2D profiles or 3D volumes becomes prohibitively large, making direct application to real datasets infeasible.[22] Therefore, in this study, ES-MDA was not applied to the entire seismic dataset but was instead implemented in a computationally feasible manner on a trace-by-trace basis to take advantage of its ability to approximate the optimal solution.

## 3. Results

### 3.1. Numerical model tests

First, a 2D layered model that included multiple complex geological features, such as geological lens bodies, faults, and pinch-outs, was used to conduct numerical tests of the proposed method. The model is shown in Figure 1, where Figure 1A–C correspond to $V_p$, $V_s$, and $\rho$, respectively. Based on Aki–Richards' approximation, the corresponding P-wave reflection coefficients were computed for four distinct incident angles: 5°, 15°, 25°, and 35°. The corresponding synthetic seismic data were generated by convolving the reflection coefficients with a 25 Hz Ricker wavelet for each incident angle. Subsequently, the synthetic seismic data were corrupted by 15% zero-mean Gaussian random noise. The contaminated data are presented in Figure 2, where Figure 2A–D correspond to incident angles of 5°, 15°, 25°, and 35°, respectively.

To simulate the incompleteness of prior information in practical applications, the true model was subjected strong smoothing process, retaining only the fundamental low-frequency background trend while completely removing the high-frequency details such as reservoirs. The smooth models are shown in Figure 3, where Figure 3A–C correspond to $V_p$, $V_s$, and $\rho$, respectively. These models were used as the a priori model.

An ensemble smoother with multiple data assimilation was then applied to the noise-contaminated seismic data. For ES-MDA, the key parameters were set as follows: the number of integrated members was 1,000, the number of inner assimilation times was 4, and the perturbation parameter for the data was 0.0001. Due to the prohibitive computational cost of applying ES-MDA to the entire 2D model, the analysis was performed on a trace-by-trace basis instead. This approach is computationally feasible and takes advantage of the ability of ES-MDA to provide a good approximation of the optimal solution. The results of ES-MDA are shown in Figure 4. It can be observed that the lateral continuity is very poor, which is expected because ES-MDA was performed trace-by-trace without lateral constraints. However, the detailed features of the model were well estimated. Therefore, the ES-MDA results still represent a good approximation of the true model.

Next, the results from ES-MDA were used as the initial solution for the AVA inversion incorporating Cauchy regularization. The corresponding inversion results are shown in Figure 5. As a baseline for comparison, traditional AVA inversion was directly applied to the noise-contaminated data using the smooth model as the initial solution. The inversion results are shown in Figure 6. Both inversion results showed good agreement with the true model. However, the inversion results in Figure 5 exhibited more detailed features, as highlighted by the black rectangular boxes. Enlarged views of these regions are shown in Figures 7 and 8. As shown in Figure 7, the lateral continuity within the lens improved, and the pinch-out locations of the lens were clearer.

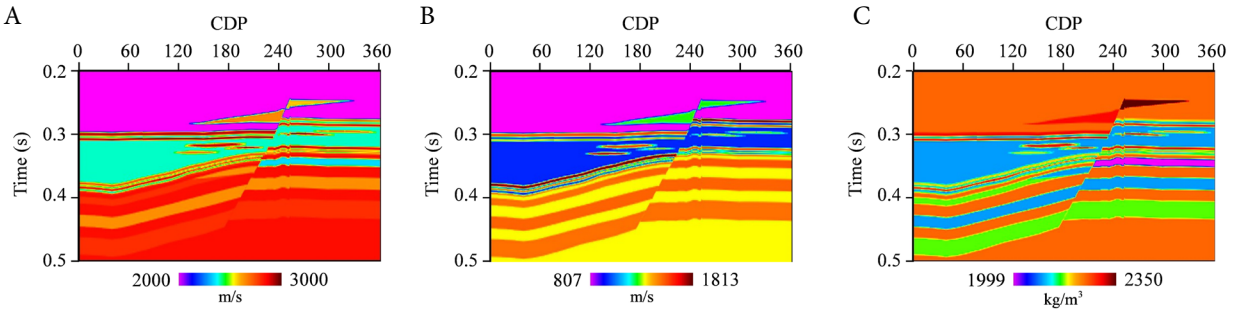



**Figure 1.** True model. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

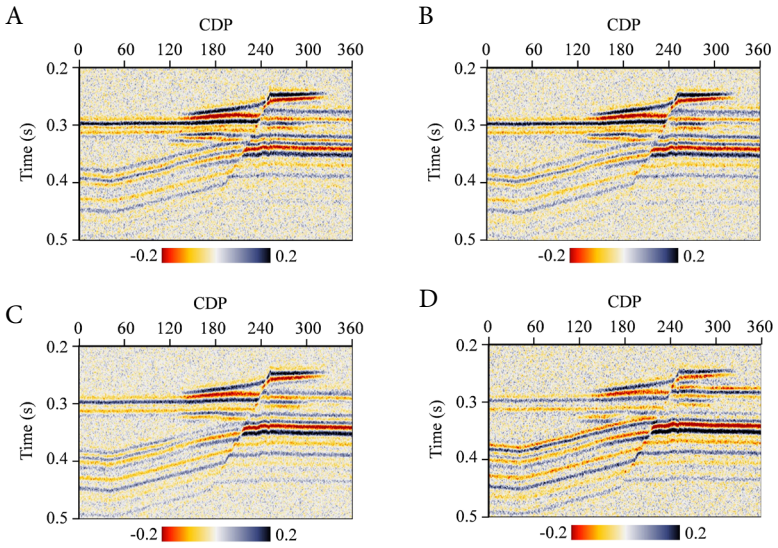



**Figure 2.** The noise-contaminated seismic data with different angles. (A) 5°, (B) 15°, (C) 25°, and (D) 35°.
Abbreviation: CDP: Common depth point.

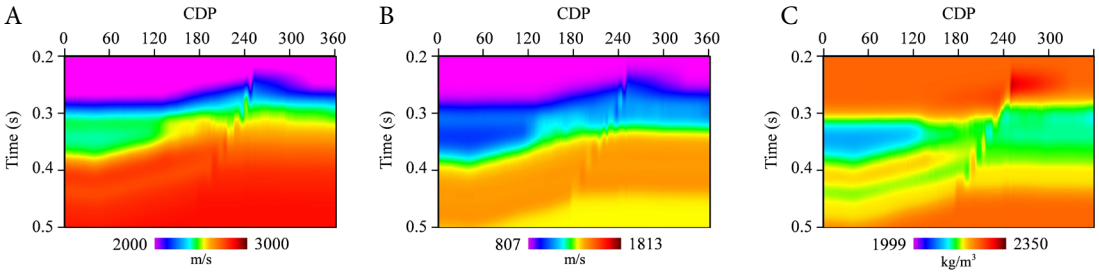



**Figure 3.** Smooth model. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

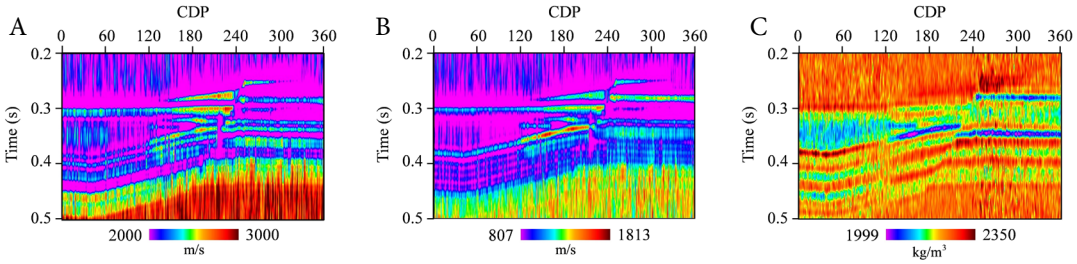



**Figure 4.** The results of ES-MDA. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; ES-MDA: Ensemble smoother with multiple data assimilation; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

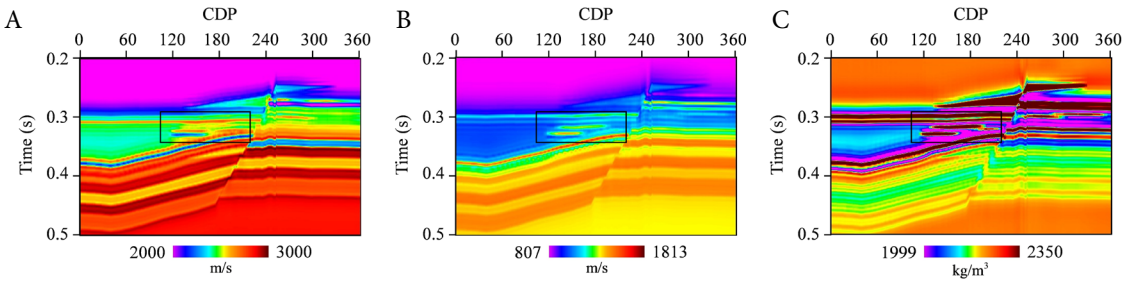



**Figure 5.** The inversion results of AVA inversion with Figure 4 as the initial solution. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: AVA: Amplitude variation with incident angle inversion; CDP: Common depth point; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

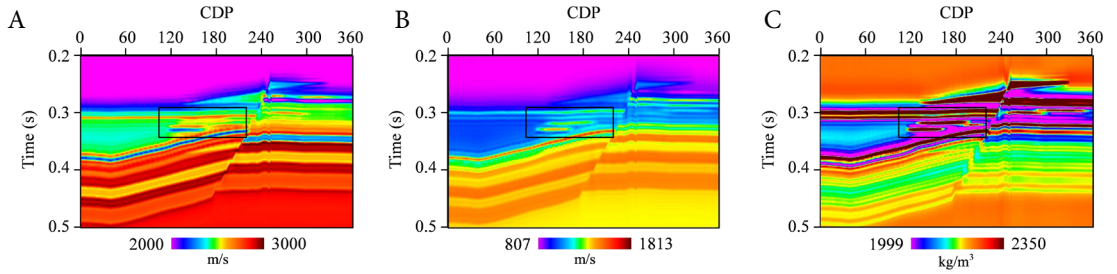



**Figure 6.** The inversion results of AVA inversion with Figure 3 as the initial solution. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

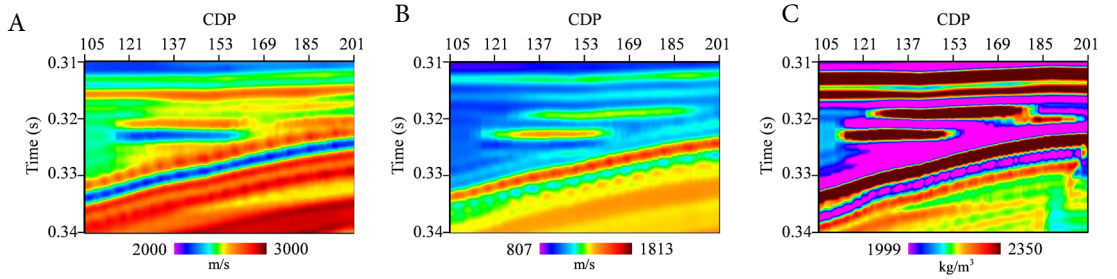



**Figure 7.** The enlarged views of the black rectangular box in Figure 5. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

Due to the relatively small differences observed in the numerical model tests between the two methods, the root mean square error (RMSE) between the inversion results and the true model was calculated for both approaches (Table 1). It can be seen that the RMSE values obtained using ES-MDA as the initial solution are smaller than those obtained using the smooth model as the initial solution.

To illustrate the stability of the inversion in the numerical model tests, the variances of the inversion results were calculated, and corresponding variance maps were obtained. Figure 9 shows the variances of the inversion result. It can be observed that, in the numerical model tests, the variance of ρ is generally larger than that of $V_p$ and $V_s$, which further confirms the well-known instability of density estimation in AVA inversion. During the quality control process, a small rectangular region within the black rectangular box in Figure 5 was used to select the optimal values of ES-MDA and regularization parameters. Consequently, for $V_p$ and $V_s$, the variances in the middle-layer region were relatively smaller. Overall, the inversion stability was considered acceptable.

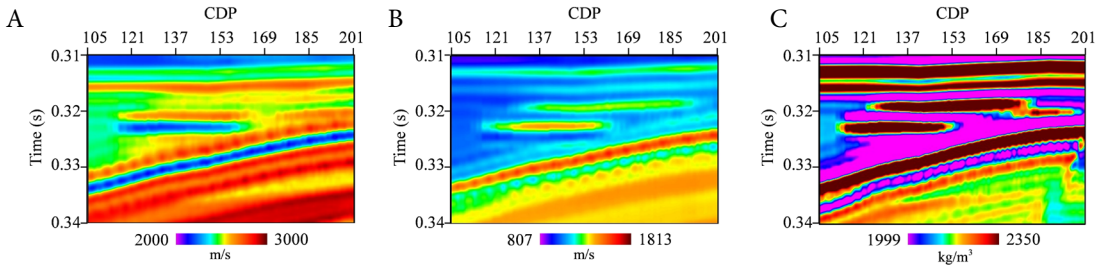



**Figure 8.** The enlarged views of the black rectangular box in Figure 6. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

**Table 1. The root mean square error between the two methods and the true model in the numerical model tests**

| Parameters | Methods | RMSE |
|---|---|---|
| P-wave velocity | AVA with a smooth model as the initial solution | 0.0513 |
| | AVA with ES-MDA as the initial solution | 0.0500 |
| S-wave velocity | AVA with a smooth model as the initial solution | 0.0784 |
| | AVA with ES-MDA as the initial solution | 0.0721 |
| Density | AVA with a smooth model as the initial solution | 0.0553 |
| | AVA with ES-MDA as the initial solution | 0.0500 |

Abbreviations: AVA: Amplitude variation with incident angle inversion; CDP: Common depth point; ES-MDA: Ensemble smoother with multiple data assimilation; RMSE: Root mean square error.

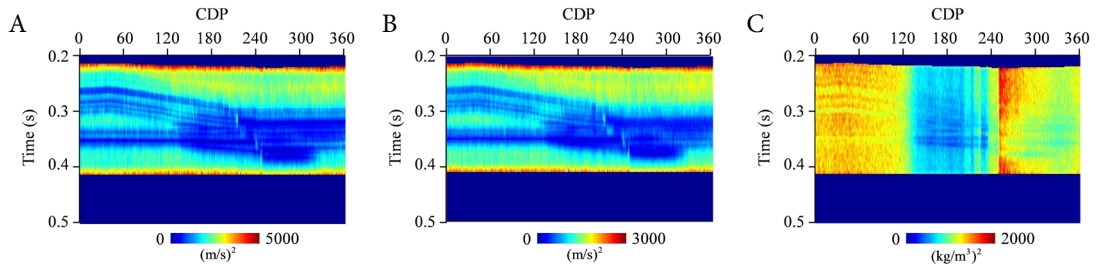



**Figure 9.** The variances of the inversion result. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

### 3.2. Real data applications

To verify the effectiveness of AVA inversion using ES-MDA results as the initial solution, the proposed method was evaluated on a real seismic line acquired from a work area in East China (exact location withheld at the data owner's request). The dataset consisted of four partial-angle stacks with angle ranges of 2–12°, 12–22°, 22–32°, and 32–42°, respectively. These sections are shown in Figure 10. Three low-frequency trend models of $V_p$, $V_s$, and $\rho$ were constructed based on well-log data and interpreted geological horizons. The dominant frequency of the seismic data is 22 Hz. Accordingly, a 22 Hz wavelet, extracted from well-log data and near-well seismic traces, was used for forward modeling in the ES-MDA process. For the ES-MDA, the key parameters were set as follows: the number of random simulations was 100, the number of inner assimilation times was 4, and the random perturbation parameter for seismic data was 0.1. The resulting ES-MDA outputs are shown in Figure 11. It can be observed that the ES-MDA results exhibited poor lateral continuity, but many detailed features were successfully reconstructed, which is critical for improving the subsequent AVA inversion results. The ES-MDA results were then used as the initial solution for AVA inversion with Cauchy regularization. For comparison, the AVA

inversion using the low-frequency trend models as the initial solution was also applied to the real seismic data. The inverted $V_p$, $V_s$, and $\rho$ obtained using the two different methods are shown in Figures 12–14, respectively. In the inverted elastic parameters profiles, the actual impedance log intersecting this seismic line was also overlaid.

From the comparison, it can be seen that the geologic structures in the inversion results from both methods are consistent with the real seismic data profiles. All inverted elastic parameter profiles showed good agreement with the actual well-log data, exhibiting sharp boundaries between upper and lower layers and demonstrating high resolution. However, the AVA inversion results initialized with the ES-MDA solution showed noticeably higher quality than those obtained using the low-frequency trend model.

In this seismic line, the target layer was located approximately between 2.42 seconds and 2.74 seconds and consisted of interbedded sandstone and mudstone layers of the river deposit. The riverbed sand bodies are stacked vertically and migrate laterally. In the western part of the seismic line, two major faults divide the target layer into three sections. The interfaces between the interbedded sandstone and mudstone layers, and the positions of the fault planes, are crucial indicators for reservoir characterization. As indicated by the arrows in Figures 12–14, the resolution of inverted elastic parameters obtained using AVA inversion with the ES-MDA initial solution improved. In addition, as indicated by the ovals in Figure 12, the inverted $V_p$ matched the well-log curve more accurately. In contrast, for AVA inversion initialized with the low-frequency trend model, the interfaces between the interbedded sandstone and mudstone layers were less distinct, and the fault break points appeared blurred. Overall, AVA inversion using the ES-MDA initial solution was better and can clearly delineate the interfaces between interbedded sandstone and mudstone layers, as well as the locations of the fault break point, compared with AVA inversion using the low-frequency trend models. Based on the ES-MDA results, which were used as the initial solution for AVA inversion, the resolution was higher than that achieved with the conventional low-frequency prior model. Consequently, the resulting AVA inversion exhibited higher quality and better agreement with the actual well-log data. Many detailed features recovered by ES-MDA were incorporated into the AVA inversion results.

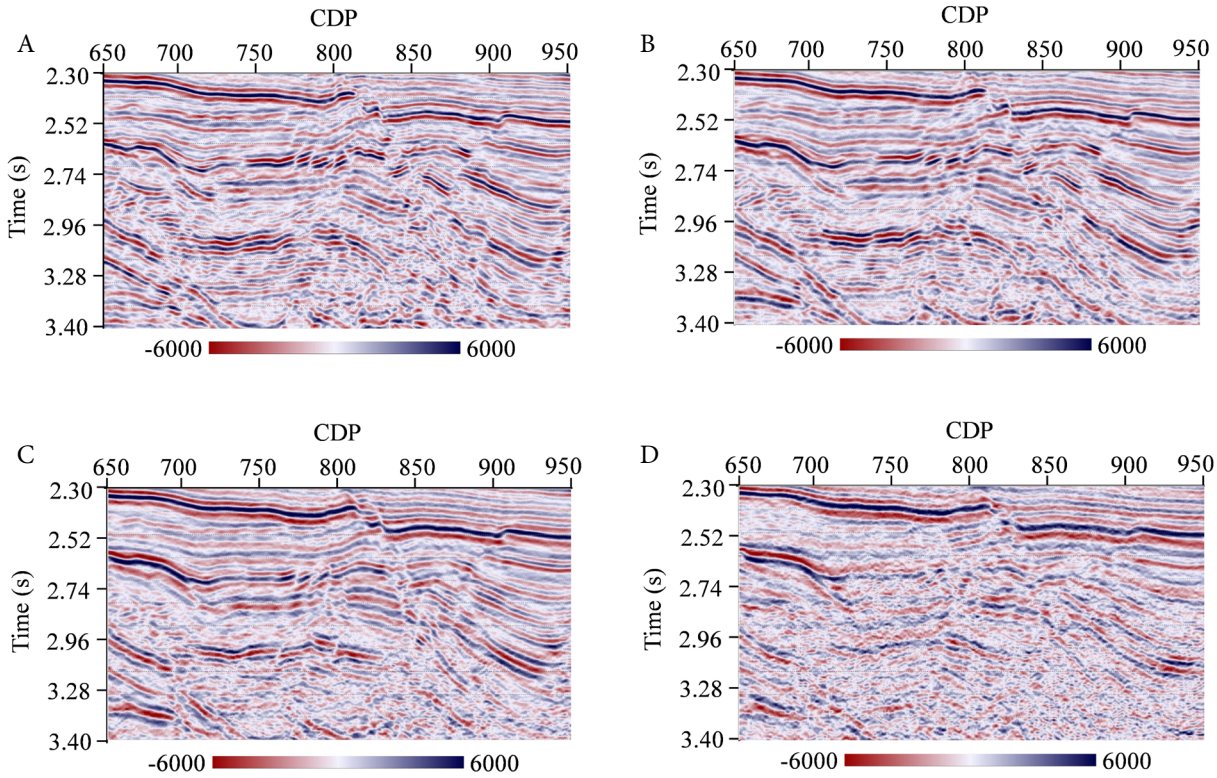



**Figure 10.** The seismic data sections with four different partial angles. (A) 2–12°; (B) 12–22°; (C) 22–32°; and (D) 32–42°.
Abbreviation: CDP: Common depth point.

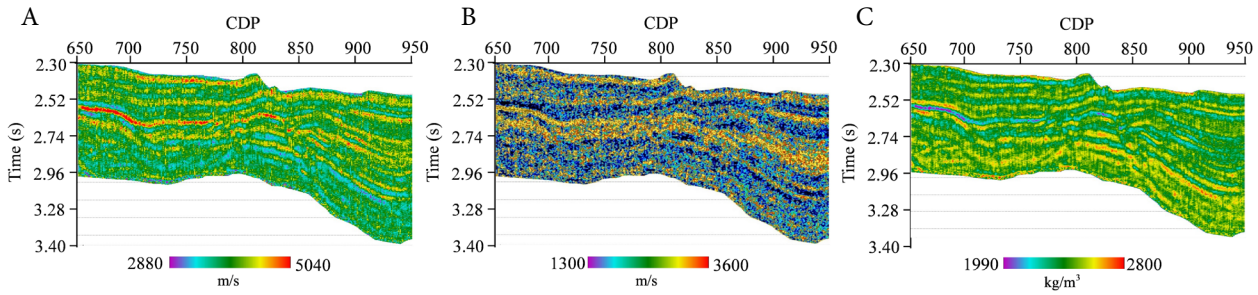



**Figure 11.** The results of ES-MDA. (A) $V_p$. (B) $V_s$. (C) $\rho$.
Abbreviations: CDP: Common depth point; ES-MDA: Ensemble smoother with multiple data assimilation; $\rho$: Density; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

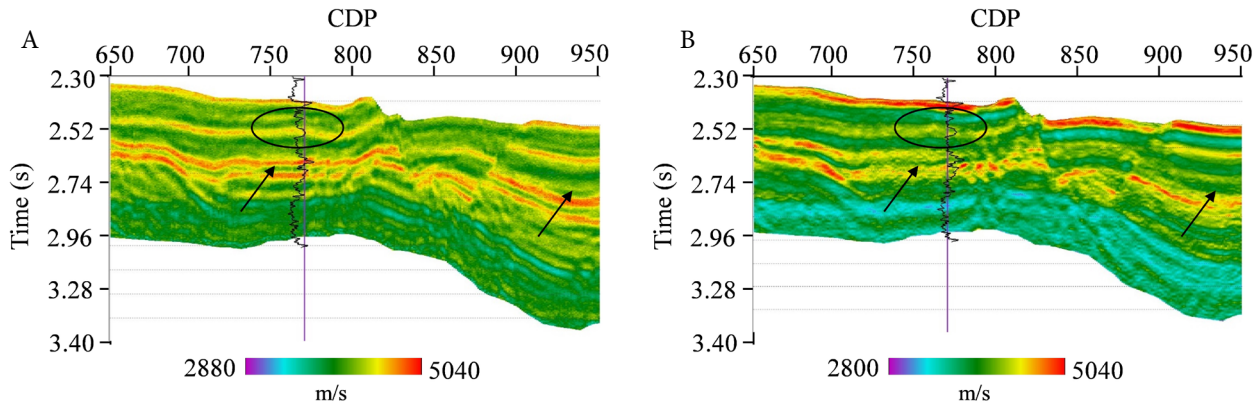



**Figure 12.** The inverted $V_p$. (A) AVA inversion with ES-MDA solution. (B) AVA inversion with a low-frequency trend model.
Abbreviations: AVA: Amplitude variation with incident angle inversion; CDP: Common depth point; ES-MDA: Ensemble smoother with multiple data assimilation; $V_p$: P-wave velocity.

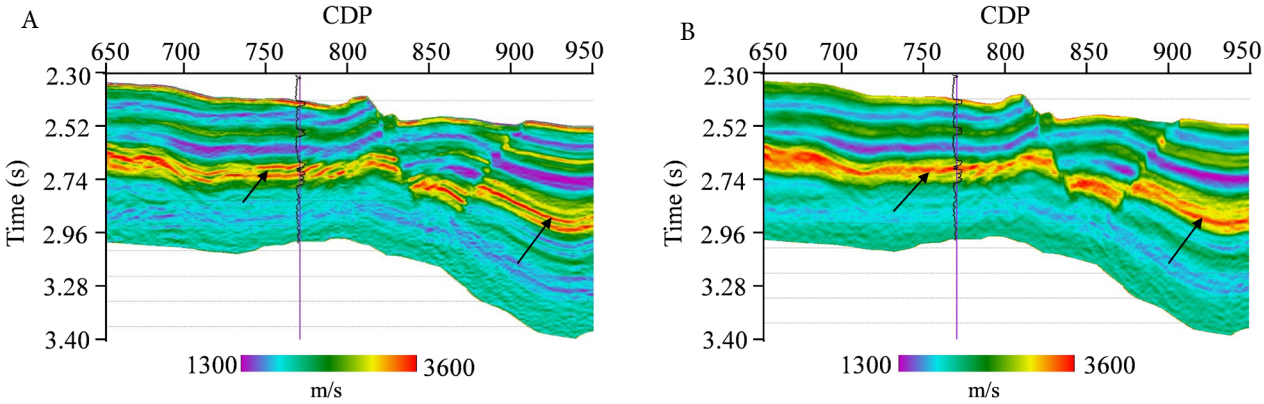



**Figure 13.** The inverted $V_s$. (A) AVA inversion with ES-MDA solution. (B) AVA inversion with a low-frequency trend model.
Abbreviations: AVA: Amplitude variation with incident angle inversion; CDP: Common depth point; ES-MDA: Ensemble smoother with multiple data assimilation; $V_s$: S-wave velocity.

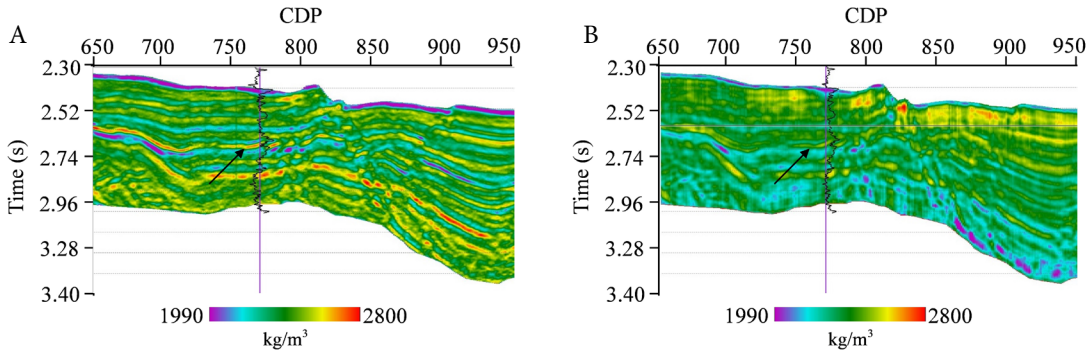



**Figure 14.** The inverted $\rho$. (A) AVA inversion with ES-MDA solution. (B) AVA inversion with a low-frequency trend model.
Abbreviations: AVA: Amplitude variation with incident angle inversion; CDP: Common depth point; ES-MDA: Ensemble smoother with multiple data assimilation; $\rho$: Density.

## 4. Discussion

In many seismic inversion methods, the objective function is non-linear due to a non-linear forward operator or the regularization term. When gradient-based algorithms are adopted, the choice of the initial model is extremely crucial. If the initial model is far from the optimal solution, the optimization may become trapped in a local minimum, and the inversion methods will yield only sub-optimal solutions. In contrast, stochastic optimal algorithms can often achieve the optimal or near-optimal solution; however, their computational cost becomes prohibitively high when dealing with large-dimensional geophysical data, as is typical in seismic reservoir characterization. Table 2 summarizes the computation times for AVA inversion, ES-MDA inversion, and AVA inversion using ES-MDA results as the initial solution in the numerical model tests. It can be seen that the computation times for AVA with a smooth initial model and AVA with ES-MDA as the initial solution are approximately equivalent. The majority of the computational cost of the proposed method in this paper is devoted to applying ES-MDA. Consequently, the number of high-dimensional ensemble models for the whole 2D profiles or 3D volumes becomes prohibitively large, making direct application to real datasets impractical. Simulating complete 2D profiles or 3D volumes necessitates the consideration of both vertical and lateral constraints. Additionally, this approach requires an extremely large amount of computation, likely exceeding the capabilities of a personal computer, and the real data would be extensive. The novelty of this work lies in combining the strengths of gradient-based algorithms and stochastic optimal algorithms while mitigating their respective limitations. Specifically, ES-MDA was first applied in a computationally feasible, trace-by-trace manner to take advantage of its ability to approximate the optimal solution. The resulting ES-MDA output was then used as a high-quality initial model for IRLS-based AVA inversion. This approach improves the accuracy of AVA inversion due to the high-quality nature of the ES-MDA solution.

In the current ES-MDA implementation, no lateral constraints were applied. Consequently, the lateral continuity of the ES-MDA solution was very poor, which somewhat limited its approximation accuracy. Considering computational efficiency, it would be beneficial to incorporate a degree of lateral constraints when simulating each trace to further improve the ES-MDA accuracy. For example, the similarity of adjacent seismic traces could be considered when drawing samples from the prior distribution of model parameters to form the initial ensemble using stochastic simulation (i.e., the solutions of adjacent seismic traces could serve as the reference for simulating the current trace. This approach could be explored in future work.

**Table 2. Computation times of the methods**

| Methods | Time (s) |
|---|---|
| AVA with a smooth model as the initial solution | 1,611.8830 |
| ES-MDA | 27,857 |
| AVA with ES-MDA as the initial solution | 1,572.9160 |

Abbreviations: AVA: Amplitude variation with incident angle inversion; ES-MDA: Ensemble smoother with multiple data assimilation.

## 5. Conclusion

In AVA inversion, Cauchy sparse regularization is frequently used to constrain the inverse problem and obtain physically meaningful results. The non-linear nature

of Cauchy regularization implies a strong dependence of the inversion quality on the initial model. In complex geological settings, the initial solution must be close to the optimal solution to ensure accurate inversion results. In this study, we successfully developed and validated an AVA inversion method that combines ES-MDA with Cauchy regularization. Specifically, the ES-MDA solution was used as the initial model for AVA inversion with Cauchy regularization. ES-MDA is a stochastic approach that integrates collective thinking with a multi-data assimilation mechanism. Its solution provides a good approximation of the optimal solution, helping to overcome the limitations of AVA inversion related to the choice of initial model.

Both numerical simulations and real-world case studies demonstrated that the proposed method improves the accuracy of estimated elastic parameters compared to AVA inversion with a smooth initial model. The research findings offer new technical approaches for seismic prediction with AVA inversion in complex reservoirs, holding practical application value. Seismic reservoir characterization often involves geophysical data of very high dimensionality, making the number of high-dimensional ensemble models for the full 2D profiles or 3D volumes prohibitively large. To address this, ES-MDA was not applied to the full seismic data profile; instead, it is performed in a computationally feasible, trace-by-trace manner, leveraging its ability to approximate the optimal solution without requiring full–scale high–dimensional simulations.

## Acknowledgments

The researchers extend their gratitude to Sinopec Shengli Oil & Gasfield Company for providing the data and technical assistance throughout this study.

## Funding

This work was supported by the National Natural Science Foundation of China under grant number: 42204123.

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Ronghuo Dai
*Formal analysis:* Rui Liao
*Investigation:* Rui Liao
*Methodology:* Ronghuo Dai
*Writing–original draft:* Rui Liao
*Writing–review & editing:* Ronghuo Dai

## Availability of data

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## References

1. Aki K, Richards PG. *Quantitative Seismology: Theory and Methods*. Vols 1-2. San Francisco, CA: Freeman; 1980:160-195.
2. Zoeppritz K, Erdbebenwellen VB. On the reflection and propagation of seismic waves. *Gottinger Nachrichten*. 1919;1:66-84.
3. Shuey R. A simplification of the Zoeppritz equations. *Geophysics*. 1985;50(4):609-614.
   doi: 10.1190/1.1441936
4. Russell BH, Gray D, Hampson DP. Linearized AVO and poroelasticity. *Geophysics*. 2011;76(3):C19-C29.
   doi: 10.1190/1.3555082
5. Xiao S, Ba J, Guo Q, Carcione JM, Zhang L, Luo C. Seismic pre-stack AVA inversion scheme based on lithology constraints. *J Geophys Eng*. 2020;17(3):411-428.
   doi: 10.1093/jge/gxaa001
6. Bao Y, Chen J, Liu XB, Zhao ZC. Joint PP and PS anisotropic AVO inversion using the exact Zoeppritz equations. *J Seism Explor*. 2021;30(6):529-544.
7. Ye T, Li J, Ding W, Long F, Yang J, Liu C. Application of full-angle prestack density inversion for deep tidal-flat thin dolomite reservoirs. *Geophys Prospect Pet*. 2024;63(6):1203-1213. [In Chinese].
   doi: 10.12431/issn.1000-1441.2024.63.06.011
8. Zhou L, Li J, Chen X, Liu X, Chen L. Prestack amplitude versus angle inversion for Young's modulus and Poisson's ratio based on the exact Zoeppritz equations. *Geophys Prospect*. 2017;65(6):1462-1476.
   doi: 10.1111/1365-2478.12493
9. Li H. Prestack seismic prediction technique for ultra-deep carbonate reservoirs in Shunbei field. *Geophys Prospect Pet*. 2025;64(4):736-748. [In Chinese].
   doi: 10.12431/issn.1000-1441.2025.0027
10. Lehocki I, Mukerji T, Avseth P, Jensen EH. Algorithms for extraction of reliable density ratios from pre-stack seismic data—Part 1: Theory. *Geophys Prospect*. 2025;73(6):e70029.
    doi: 10.1111/1365-2478.70029
11. Sun W, Chen Z, Wang R, Duan M. Research on fluid identification technology of submarine fan reservoirs: A case study of Meishan Formation in Ledong–Lingshui Sag. *Geophys Prospect Pet*. 2025;64(6):1107-1117. [In Chinese].
    doi: 10.12431/issn.1000-1441.2024.0230
12. 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 Seism Explor*. 2025;34(5):1-17.

doi: 10.36922/JSE025310048

13. Sacchi MD. Reweighting strategies in seismic deconvolution. *Geophys J Int*. 1997;129(3):651-656.

doi: 10.1111/j.1365-246X.1997.tb04500.x

14. Alemie W, Sacchi MD. High-resolution three-term AVO inversion by means of a Trivariate Cauchy probability distribution. *Geophysics*. 2011;76(3):R43-R55.

doi: 10.1190/1.3554627

15. Zhang F, Dai R. Nonlinear inversion of pre-stack seismic data using variable metric method. *J Appl Geophys*. 2016;129:111-125.

doi: 10.1016/j.jappgeo.2016.03.035

16. Dai R, Yin C. An iterative re-weighting algorithm for element to solve sparsity-regularized linear inverse problem and an application in sparse-spike deconvolution of seismic data. *IEEE Geosci Remote Sens Lett*. 2025;22:7506605.

doi: 10.1109/LGRS.2025.3574446

17. Zhang R, Castagna J. Seismic sparse-layer reflectivity inversion using basis pursuit decomposition. *Geophysics*. 2011;76(6):R147-R158.

doi: 10.1190/geo2011-0103.1

18. Zong Z, Yin X, Wu G. Geofluid discrimination incorporating poroelasticity and seismic reflection inversion. *Surv Geophys*. 2015;36(5):659-681.

doi: 10.1007/s10712-015-9330-6

19. Wang Y. Basics of seismic inversion. In: *Seismic Inversion: Theory and Applications*. John Wiley & Sons; 2016:45-80.

doi: 10.1002/9781119258032.ch1

20. Emerick AA, Reynolds AC. Ensemble smoother with multiple data assimilation. *Comput Geosci*. 2013;55:3-15.

doi: 10.1016/j.cageo.2012.03.011

21. Gineste M, Eidsvik J. Seismic waveform inversion using the ensemble Kalman smoother. In: Proceedings of the *79th EAGE conference and exhibition 2017*; June 12-15, 2017; Paris, France. European Association of Geoscientists & Engineers; 2017:1-5.

doi: 10.3997/2214-4609.201700794

22. Grana D, Mukerji T, Doyen P. *Seismic Reservoir Modeling: Theory, Examples, and Algorithms*. John Wiley & Sons; 2021:130-145. Available from: https://www.wiley.com/en-us/Seismic+Reservoir+Modeling%3A+Theory%2C+Examples%2C+and+Algorithms-p-9781119086185 [Last accessed in December 2025].

23. Li C, Zhang F. Amplitude-versus-angle inversion based on the l1-norm-based likelihood function and the total variation regularization constraint. *Geophysics*. 2017;82(3):R173-R182.

doi: 10.1190/geo2016-0182.1

24. Dai R, Yang J. Amplitude-versus-angle (AVA) inversion for pre-stack seismic data with L0-Norm-Gradient regularization. *Mathematics*. 2023;11(4):880.

doi: 10.3390/math11040880

25. Fatti JL, Smith GC, Vail PJ, Strauss PJ, Levitt PR. Detection of gas in sandstone reservoirs using AVO analysis: A 3-D seismic case history using the Geostack technique. *Geophysics*. 1994;59(9):1362-1376.

doi: 10.1190/1.1443695

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Advances in intelligent seismic interpretation for onshore unconventional reservoirs in CNOOC: The Linxing–Shenfu gas field case

**Wenlan Li[1], Qixin Li[2], Xiaowen Zheng[3], Bo Wang[1]*, Di Wang[2], and Jicai Ding[2]**

[1]China United Coalbed Methane Corporation Ltd., Beijing, China

[2]China National Offshore Oil Corporation Research Institute Corporation Ltd., Beijing, China

[3]China National Offshore Oil Corporation Gas And Power Group Corporation Ltd., Beijing, China

(This article belongs to the *Special Issue: Geophysical Inversion and Intelligent Prediction Technologies for Complex Hydrocarbon Reservoirs*)

## Abstract

The Linxing–Shenfu gas field, a key block for China National Offshore Oil Corporation's onshore unconventional oil and gas exploration, is characterized by complex geology that poses dual challenges to the efficiency and accuracy of traditional seismic interpretation methods. This study presents a systematic review of the application progress of machine learning, particularly deep learning, in seismic interpretation within this block since 2018. To address the specific geological characteristics and exploration needs, we developed a comprehensive intelligent interpretation workflow. This workflow integrates intelligent horizon and fault interpretation, deep clustering for seismic facies analysis, and automated identification of special geological bodies (e.g., Zijinshan igneous rock mass), enabling the accurate reconstruction of the stratigraphic framework. Furthermore, leveraging deep learning models, we achieved direct prediction of lithology, physical properties (e.g., porosity, permeability), and gas-bearing parameters, culminating in the comprehensive characterization of geological "sweet spots." Practical applications demonstrate that this intelligent interpretation workflow not only significantly enhances interpretation efficiency but also provides distinct advantages for overcoming the bottlenecks of traditional theoretical methods, such as handling low signal-to-noise ratio data, identifying thin interbeds, and predicting "sweet spots." This review provides robust support for efficient exploration and development decision-making in the Linxing–Shenfu Block.

***Keywords:*** Unconventional reservoir; Seismic interpretation; Machine learning; Deep learning; Sweet spot

***Corresponding author:**
Bo Wang
(wangbo64@cnooc.com.cn)

**Citation:** Li W, Li Q, Zheng X, Wang B, Wang D, Ding J. Advances in intelligent seismic interpretation for onshore unconventional reservoirs in CNOOC: The Linxing–Shenfu gas field case. *J Seismic Explor.* 2026;35(2):025460108. doi: 10.36922/JSE025460108

**Received:** November 11, 2025

**Revised:** February 3, 2026

**Accepted:** February 7, 2026

**Published online:** March 31, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

The Linxing–Shenfu Block of China National Offshore Oil Corporation (CNOOC), located in the northeastern part of the Ordos Basin, covers an area of approximately 4,500 km$^2$. The exploration history of the Linxing–Shenfu Block comprises three

distinct phases: (i) the discovery of peripheral gas fields, (ii) initial gas field exploration, and (iii) large-scale gas field development. In 2012, CNOOC entered the Linxing–Shenfu Block after systematically comparing geological conditions between the basin margin and the interior—a strategic shift toward tight gas exploration was proposed. From 2014 to 2015, significant exploration breakthroughs were achieved in the L-4 well area, with high-yield gas flows from the Taiyuan formation and proven geological reserves exceeding 20 billion $m^3$, which marked the successful initial development of the tight gas reservoir.

Through a series of exploration activities, including drilling, seismic acquisition, logging evaluation, and fracture testing, geological understanding and technological innovation were deepened, gradually revealing the geological characteristics of the gas field, and resulting in rapid reserve growth. This led to the discovery of the Linxing tight gas field with a capacity of over 100 billion $m^3$ in 2019. Subsequently, with breakthroughs in deep coalbed methane (CBM) exploration, the exploration strategy for Linxing–Shenfu Block shifted to a dual focus on both tight gas and deep CBM,[1] with cumulative proven geological reserves of 280 billion $m^3$ of tight gas and 270 billion $m^3$ of deep CBM. In recent years, under the guidance of the whole hydrocarbon system theory in coal measures, exploration research has gradually increased, focusing on the Ordovician carbonate rocks of the Lower Paleozoic, Carboniferous bauxite rocks, and the Zijinshan igneous rock mass. These efforts are expected to pave the way for new replacement areas for future energy development.

Since 2013, through integrated planning and phased implementation, a cumulative total of 3,500 $km^2$ of three-dimensional (3D) seismic data has been acquired in the Linxing–Shenfu Block, achieving near-complete coverage of the exploration area. The application of 3D seismic technology has played an important role in the discovery of tight gas fields and CBM fields, as well as in the preliminary exploration of new fields such as bauxite, carbonate rocks, and igneous rocks. In terms of 3D seismic interpretation techniques, a stratigraphically facies-controlled tight-sandstone gas prediction technology system has been developed to address seismic forecasting requirements for geological and engineering sweet spots.[2] Techniques such as pre-stack waveform-controlled high-resolution inversion, geostatistical coal facies prediction, and coal body structure prediction based on multi-attribute clustering have also been explored.

Since 2018, machine learning (ML) technology, especially deep learning (DL), has seen explosive growth in cross-disciplinary research with geophysics,[3,4] and data-driven geophysical interpretation techniques have begun to gain attention. Meanwhile, CNOOC's onshore unconventional oil and gas exploration has accumulated substantial datasets characterized by diversity, homogeneity, and density. Under the guidance of new technologies and robust data support, CNOOC has developed a comprehensive ML-based seismic interpretation framework for onshore unconventional oil and gas over six years of sustained research, including horizon and fault interpretation, seismic facies clustering, special rock body characterization, and lithological, petrophysical, and gas-bearing property prediction. It encompasses the complete seismic interpretation workflow, including the establishment of stratigraphic frameworks, lithological assemblage classification, target lithology prediction, and the prediction of physical and gas-bearing properties. The combination of data-driven ML seismic interpretation technology and model-driven conventional seismic interpretation technology has achieved good results in the exploration and development of tight sandstone gas and deep CBM in the Linxing–Shenfu Block, as well as in the exploration of Ordovician carbonate rocks, Carboniferous bauxite rocks, and the Zijinshan igneous body.

## 2. Geological conditions and the demand for intelligent interpretation

### 2.1. Geological characteristics of the Linxing–Shenfu Block

The Linxing–Shenfu Block resides within the transitional zone between the Yishan Slope and the Jinxi flexural fold belt in the Ordos Basin (west of the Lishi Fault Zone), and exhibits structural characteristics comparable to those of the Yishan Slope (Figure 1).[5] Within the block, Upper Paleozoic strata exhibit an unconformable contact with Lower Paleozoic strata, marked by the absence of Middle-Upper Ordovician Series, Silurian System, Devonian System, and Lower Carboniferous Series. According to current drilling data, the stratigraphic sequence from bottom to top consists of the following: the Lower Ordovician Majiagou Formation, and the Upper Paleozoic strata exhibit continuous deposition with conformable contacts, dominated by transitional marine-terrestrial facies and continental clastic rock deposits. From base to top, the sequence includes the (i) Upper Carboniferous Benxi Formation, the Lower Permian Taiyuan and Shanxi Formations, the Middle Permian Shihezi Formation, the Upper Permian Shiqianfeng Formation, and the Lower Triassic Liujiagou Formation; (ii) the Upper Pliocene of the Neogene; (iii) the Upper Pleistocene of the Quaternary; (iv) and the Holocene of the Quaternary.

Based on the spatial relationship between the reservoirs and the source rock, the Linxing–Shenfu Block can be

divided into three sets of oil and gas systems in the vertical direction (Figure 2): (i) tight gas with a lower source and upper reservoir configuration, (ii) self-sourced CBM gas, and (iii) multi-sourced gas in carbonate and bauxite rocks with multiple reservoirs.[6,7] In addition, influenced by the thermal effects of Yanshanian magmatic intrusion in the eastern part of the basin, the Zijinshan igneous intrusion induced localized strata uplift in the central part of Linxing, resulting in the exposure of the Zijinshan alkaline complex. Its magmatic thermal effect significantly influences hydrocarbon expulsion from source rocks and the regional structural framework.[8] Additionally, the intrusion and its associated metamorphic system represent a key exploration target in the Linxing–Shenfu Block.

### 2.2. Demand for intelligent seismic interpretation technology

Bergen *et al.*[9] proposed three categories of ML application in geoscience: (i) automating complex predictions beyond explicit commands; (ii) modeling and inverse problems to approximate numerical simulations or capture relationships; and (iii) discovering new patterns, structures, and relationships. Based on the exploration and development demands of the Linxing–Shenfu Block, intelligent seismic interpretation technology has demonstrated significant efficacy in two aspects: labor-intensive interpretation tasks and seismic interpretation where traditional theories are inapplicable, specifically manifested in:

(i) The stratigraphic sequence interpretation across 24 geological units—including five subsections of the Ordovician Majiagou Formation, two subsections of the Carboniferous Benxi Formation, two subsections of the Taiyuan Formation, two subsections of the Shanxi Formation, eight subsections of the Shihezi Formation in the Permian, and five subsections of the Triassic Shiqianfeng Formation—has been substantially enhanced in efficiency, precision, and consistency through intelligent horizon interpretation techniques.

(ii) For the two sets of fault systems in the Upper and Lower Paleozoic, intelligent fault interpretation has been realized with the assistance of automated fault surface extraction tools, trained on a limited set of manually interpreted labels.

(iii) The essence of intelligent interpretation of stratigraphic horizons and faults is the intelligent construction of a stratigraphic framework. Within this framework, spatial segmentation and clustering of lithological assemblages across stratigraphic units are critical. Given the low signal-to-noise ratio (SNR) seismic data in the Linxing–Shenfu Block, intelligent seismic facies analysis is the most effective method for automated feature extraction.

(iv) Beyond typical sedimentary strata, the Zijinshan igneous intrusion in the eastern part of the Linxing Block significantly influences regional hydrocarbon source rocks and tectonic evolution. The 3D

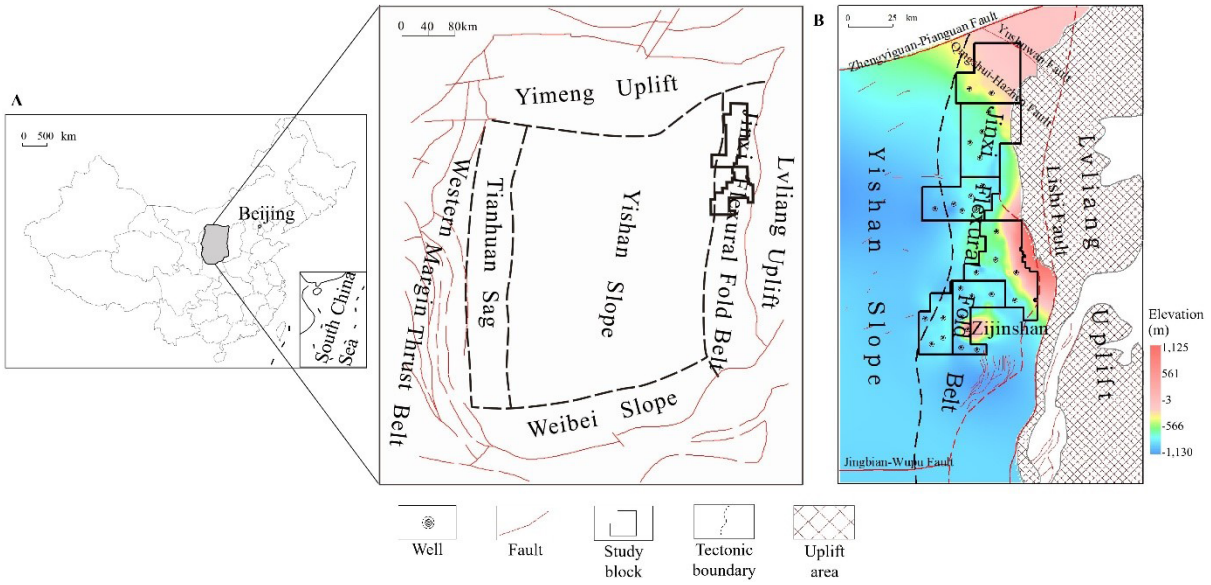



**Figure 1.** Geological overview of the study area. (A) Location of the Ordos Basin within China and its tectonic map. The Ordos Basin comprises six major geological units: the Jinxi flexural fold belt, the Yimeng uplift, the Yishan slope, the Weibei slope, the Tianhuan sag, and the Western Margin Thrust Belt. (B) Structural configuration of the Linxing–Shenfu Block in the eastern part of the basin.

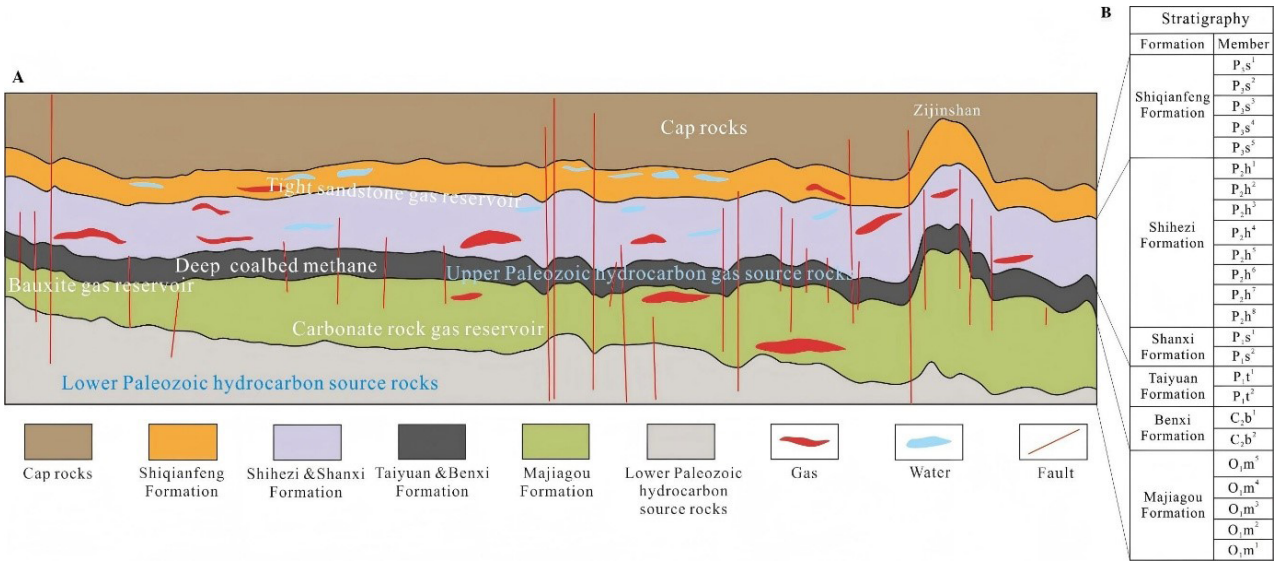



**Figure 2.** Geological model diagram. (A) Reservoir formation model of the entire oil and gas system in the Linxing–Shenfu Block. The system contains dual source rocks (Upper and Lower Paleozoic) and three major reservoir types (tight sandstone, bauxite, and carbonate gas reservoirs). (B) The stratigraphic sequence of the study area, comprising 24 defined geological units.

characterization of the igneous intrusion is a complex stratigraphic interpretation process with multiple *Z*-values in the vertical direction, but intelligent identification technology for special geological bodies can reduce the complexity of interpretation.

(v) The exploration and development of tight sandstone gas and CBM in the Upper Paleozoic involve predicting geological "sweet spots" (areas with high porosity, high permeability, and high gas saturation) characterized by lithology, petrophysical properties, and gas saturation. Under low porosity and permeability conditions, the overlap of seismic elastic parameters reduces their sensitivity to lithology, petrophysical properties, and gas-bearing properties. Supported by extensive well data, data-driven intelligent prediction of these parameters enables robust support for well placement optimization and reserve estimation.

## 3. A machine learning-driven workflow for integrated seismic interpretation

### 3.1. The proposed workflow and its novelty

This study presents a fully integrated and intelligent seismic interpretation workflow that systematically progresses from raw data to the characterization of "sweet spots," demonstrating a paradigm shift toward automation in subsurface analysis (Figure 3). The workflow begins with the robust, DL-enhanced identification of fundamental geological structures under low-SNR conditions, where a convolutional neural network (CNN) combined with conditional random fields (CRF) refinement and path optimization automates horizon tracking, and a U-Net++ architecture directly detects faults from seismic data, effectively replacing manual picking and traditional attribute computation. Building upon this structural framework, the workflow automatically interprets geological meaning using a deep clustering auto-encoder for seismic facies mapping and a U-Net for semantic segmentation of specific geological bodies, such as igneous rock masses, thereby bypassing subjective manual attribute selection.

Leveraging these structural and facies constraints, the system then transitions to quantitative reservoir evaluation, employing a Markov chain-constrained Bayesian classification framework for high-resolution lithofacies discrimination in thin-bedded coal-bearing strata and a geology-informed DL model for predicting physical properties such as porosity, permeability, and gas saturation. The integrated workflow culminates in a multi-layer feed-forward neural network that integrates seismic data with corrected well logs to establish nonlinear mappings between optimized seismic attributes and corrected total gas (TG) curves, effectively resolving fluid ambiguity under complex rock physics conditions and generating a continuous gas-bearing probability volume. Ultimately, all outputs, structural model, facies, lithology, physical properties, and gas content, are fused within a unified 3D model to automatically delineate composite "sweet spots."

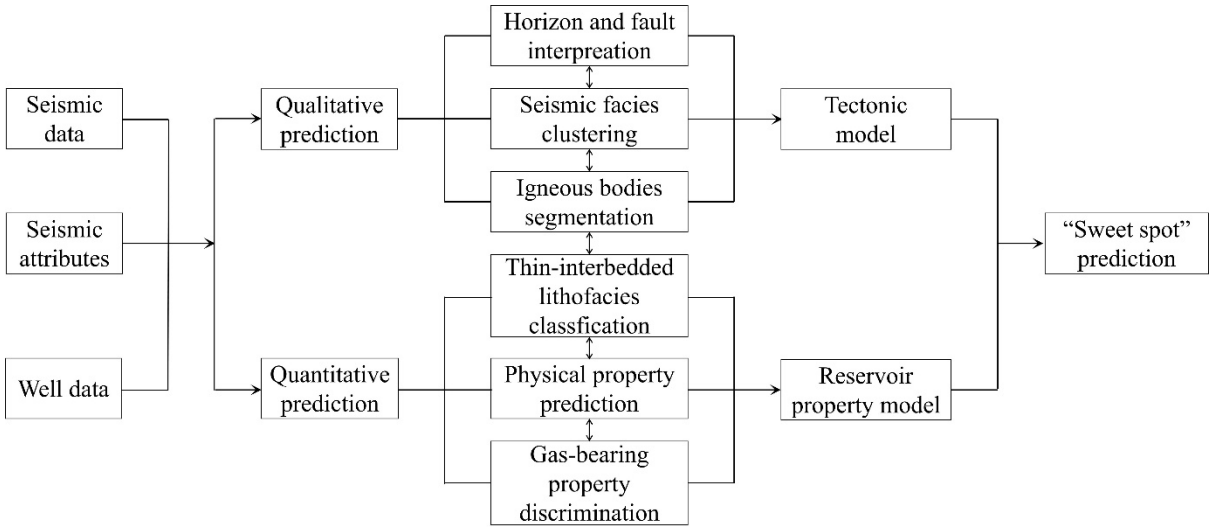



**Figure 3.** The machine learning-driven workflow for seismic interpretation

This innovative workflow, which establishes a systematic, automated solution, significantly reduces interpretive subjectivity and manual effort, thereby setting a new standard for data-driven reservoir characterization in geologically complex settings.

### 3.2. Machine learning-enhanced identification for complex geological structures

#### *3.2.1. Robust horizon and fault interpretation under low signal-to-noise ratio*

Accurate structural interpretation, encompassing both seismic horizon tracking and fault identification, is a cornerstone of subsurface model building for hydrocarbon exploration.[10] However, in the Linxing–Shenfu Block, this task is significantly impeded by low SNR and weak lateral continuity of seismic data due to heterogeneous sedimentation. Traditional horizon interpretation relies heavily on manual picking by geologists, a subjective and time-consuming process.[11] While semi-automated methods such as seed-point–guided tracking improve efficiency by correlating local waveform similarities,[12] their accuracy remains limited when encountering significant lateral waveform variations or fault disruptions. Fault interpretation traditionally relies on manual detection of seismic discontinuities, a process prone to subjectivity and constrained by data quality.[13] Although seismic attribute algorithms (e.g., coherence,[14] variance,[15] curvature,[16] and ant tracking[17]) have enhanced interpretation efficiency, they still face limitations in discriminating faults from lithological boundaries and remain computationally intensive and resolution-sensitive. This study addresses the challenge of achieving robust, automated structural interpretation under such seismically adverse conditions.

Recent advancements in DL, particularly CNN, have demonstrated significant progress in horizon interpretation[18-20] and fault detection.[21-24] Therefore, we developed an integrated workflow combining CNNs with CRF and path-optimization algorithms, forming a three-stage processing framework that progresses from local feature extraction to preliminary interpretation and, finally, to globally optimal interpretation.

In the initial stage, CNNs automatically detect features from seismic data. For horizon interpretation, CNNs are used to learn local waveform patterns, generating an initial probability volume that highlights potential horizon locations. Concurrently, for fault interpretation, specific data conditioning is applied to enhance structural edges: a multi-step structure-oriented filter is employed to suppress random noise while preserving discontinuities. Furthermore, amplitude balancing is applied to the fault-detection input to equalize energy levels across the volume, ensuring that fault features in low-amplitude or deep zones are effectively captured. Subsequently, a U-Net++ network[25] with dense skip connections executes end-to-end fault detection, producing a 3D fault probability volume. This replaces manual feature picking and traditional attribute computation with robust, initial pattern recognition.

However, the CNN's horizon predictions, while accurate locally, lack spatial coherence due to its limited receptive field. Similarly, the raw fault probability volume

may contain noise and exhibit discontinuous fault planes. Therefore, we transformed the preliminary model into a geologically consistent interpretation. For horizons, this is achieved through a synergistic combination of CRF and a path-optimization algorithm—the CRF serves as a refinement tool, and the path-optimization algorithm performs a global search to extract the most probable continuous horizon. As a discriminative probabilistic graphical model, CRF can effectively capture state–state dependencies in sequential data (e.g., continuous strata), rather than only focusing on feature mapping. We used the CNN model's output as input to the CRF model and trained it using horizon data as labels. By learning the transition probability matrix between formations, the CRF imposes strict global constraints on the predicted sequence, thereby effectively avoiding stratigraphic sequence reversals or jumps caused by local noise. For faults, the refinement involved automated extraction of fault patches from the probability volume, followed by interactive editing and validation to ensure geological reasonableness, ultimately generating fault sticks for 3D structural modeling.

Compared to seed-point auto-tracking, our method maintains continuous and accurate horizons in challenging zones (Figure 4), demonstrating improved lateral continuity and consistency with well tops. The DL-derived faults exhibited greater clarity, better vertical continuity, and more geologically reasonable geometries than those from variance or coherence attributes (Figure 5). The resulting fault probability volume enables efficient extraction of fault patches for 3D modeling (Figure 6). In summary, this integrated approach delivers structurally superior interpretations in challenging data, significantly reducing interpretive ambiguity and manual effort.

### *3.2.2. Seismic facies clustering with deep embedding representation*

Seismic facies analysis serves as a fundamental methodology for interpreting subsurface sedimentary environments, lithological assemblages, and reservoir characteristics.[26] Traditional approaches—ranging from manual interpretation to attribute-based clustering algorithms[27] (e.g., K-means, hierarchical clustering, and

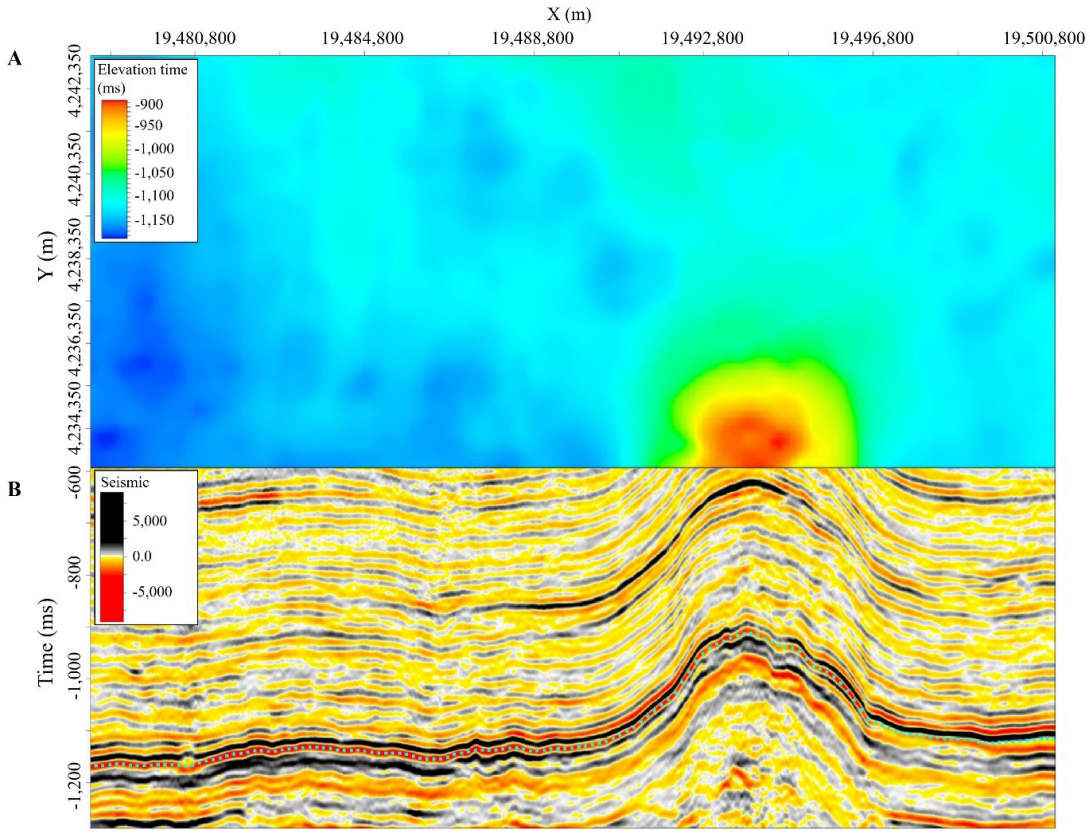



**Figure 4.** Deep learning-based horizon interpretation results for the Linxing Block. (A) Time-structure map derived from horizon interpretation. (B) Seismic profile showing the interpreted horizon, demonstrating consistency with primary seismic reflections.

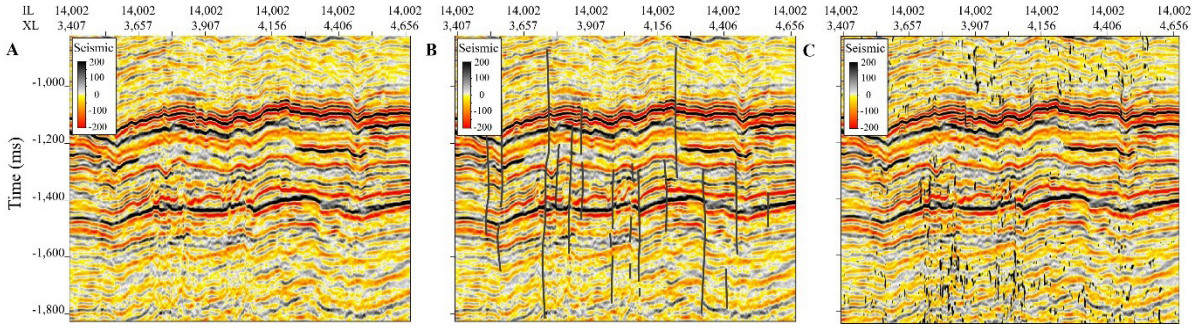



**Figure 5.** Comparison of fault detection methods on a representative seismic section. (A) Raw seismic profile without interpretation. (B) Fault planes extracted from the deep learning-based probability volume, demonstrating superior vertical continuity and geometric clarity. (C) Faults highlighted using the conventional variance attribute, appearing noisier and more discontinuous.

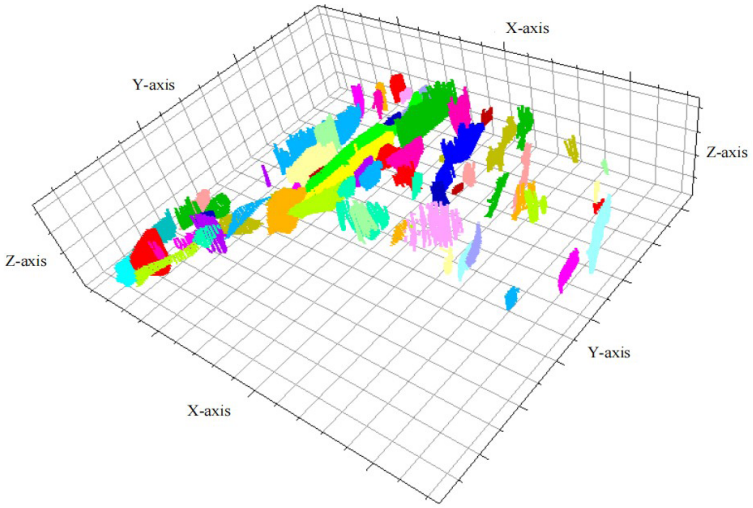



**Figure 6.** Polygons and fault sticks predicted using deep learning. Colors distinguish individual faults.

probabilistic clustering) and classification algorithms—inherently depend on manually designed seismic attributes.[28] This reliance introduces subjectivity, incompleteness, and redundancy, as the pre-selected attributes may not constitute the optimal feature space for facies discrimination.[29] With thin-interbedded reservoirs, such as tight sandstone, becoming key exploration targets, the rapid lithological heterogeneity within channels demands greater accuracy in seismic facies interpretation. Consequently, it is imperative to develop a fully data-driven framework. The core of this framework lies in its ability to skip manual attribute selection, directly extract optimal features from raw seismic waveforms to distinguish sedimentary facies, and achieve end-to-end mapping from waveforms to facies categories.

Crucially, unlike the structural task, data conditioning for facies clustering prioritizes amplitude preservation. We utilized seismic data processed through a true relative amplitude workflow, avoiding aggressive amplitude gain control or non-linear filtering that could distort lithological information. This ensures that the latent features extracted by the deep CNN reflect genuine sedimentary variations rather than processing artifacts.

Inspired by successful applications of deep CNNs in seismic facies analysis[30-33] and by classical mathematical transformations in geophysics (e.g., Fourier and wavelet transforms), we propose a DL-based seismic facies clustering method that balances clustering and representation performance.[34] Its key features lie in learning seismic attributes that not only restore the original signal non-destructively but are also suitable for seismic facies clustering. The architecture of our network (Figure 7) was designed to incorporate both a reconstruction loss function ($L_{rec}$) and a clustering loss function ($L_c$) during training by constructing and optimizing a joint loss function.

The implementation consisted of four sequential steps:

(i) The auto-encoder network was initially trained by minimizing the L2 norm between the input and output, thereby optimizing $L_{rec}$ to establish fundamental feature representation capabilities.
(ii) The trained encoder was subsequently employed to transform raw seismic data into latent representations, followed by K-means clustering to initialize the cluster centers within this hidden encoding space.
(iii) Student's *t* distribution was adopted to quantify the similarity between any point in the latent space and the cluster centers, forming the target probability distribution *P*. An auxiliary target distribution *Q* was constructed to enhance intra-cluster purity. The Kullback–Leibler divergence between these distributions served as $L_c$, which was then back-propagated jointly with $L_{rec}$ to optimize the network parameters. The hyperparameter λ in the joint loss function balanced the trade-off between reconstruction loss and clustering loss, and its value was determined empirically to ensure the learned representations retain sufficient geologic information while forming distinct clusters, with λ=0.1 chosen herein via validation. This selection aims to prevent gradient instability from a large λ and avoid poor cluster separation from a small λ, ultimately optimizing for both representation learning and clustering performance.
(iv) The optimized auto-encoder network was finally deployed across the seismic survey area, where each seismic trace is assigned a facies label based on minimum Euclidean distance to the cluster centers, generating the final planar seismic facies.

Validation in the Linxing–Shenfu Block demonstrates the superior performance of the proposed method compared to conventional seismic attributes. As illustrated in Figure 8, the seismic facies map generated by our network provided more geologically meaningful characterization than the root mean square amplitude attribute. The advantages are manifested in three key aspects:

(i) Enhanced channel delineation: While the root mean square amplitude attribute clearly delineated two main channels in the central study area, our method not only accurately captured these two channels but also revealed an additional channel (Face 6) in the western region, which has been confirmed by drilling data.
(ii) Improved inter-channel differentiation: The method successfully discriminated between different channel types. Specifically, the western channel (Face 6) in the central area exhibited distinct reservoir properties from the eastern channel (Face 4), with drilling data confirming differences in average gas layer thickness, porosity, and gas saturation between these facies.
(iii) Enriched geological details: The method revealed finer sedimentary features, including channel boundaries (Face 3) and local thickness variations within channel deposits (Face 2), effectively capturing the internal heterogeneity of the channel systems. These results validate the robust clustering capability and data recovery accuracy of our proposed network, establishing it as a more effective tool for seismic facies analysis and reservoir prediction in geologically complex settings.

### *3.2.3. Semantic segmentation of igneous intrusive bodies with limited labels*

Igneous rocks and salt-related structures play multiple roles in hydrocarbon exploration, both affecting the understanding of subsurface structures and the reconstruction of sedimentary environments,[35] and also serving as significant oil and gas reservoirs. These

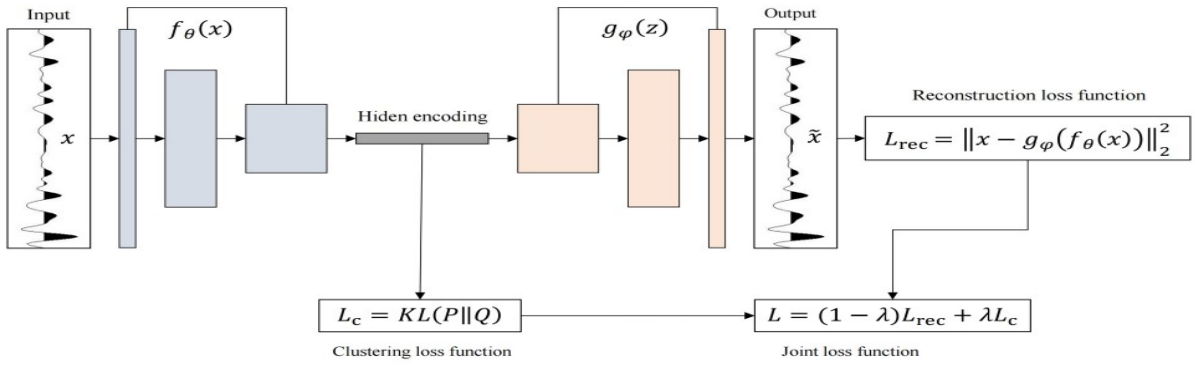



**Figure 7.** Structure of seismic facies deep clustering network

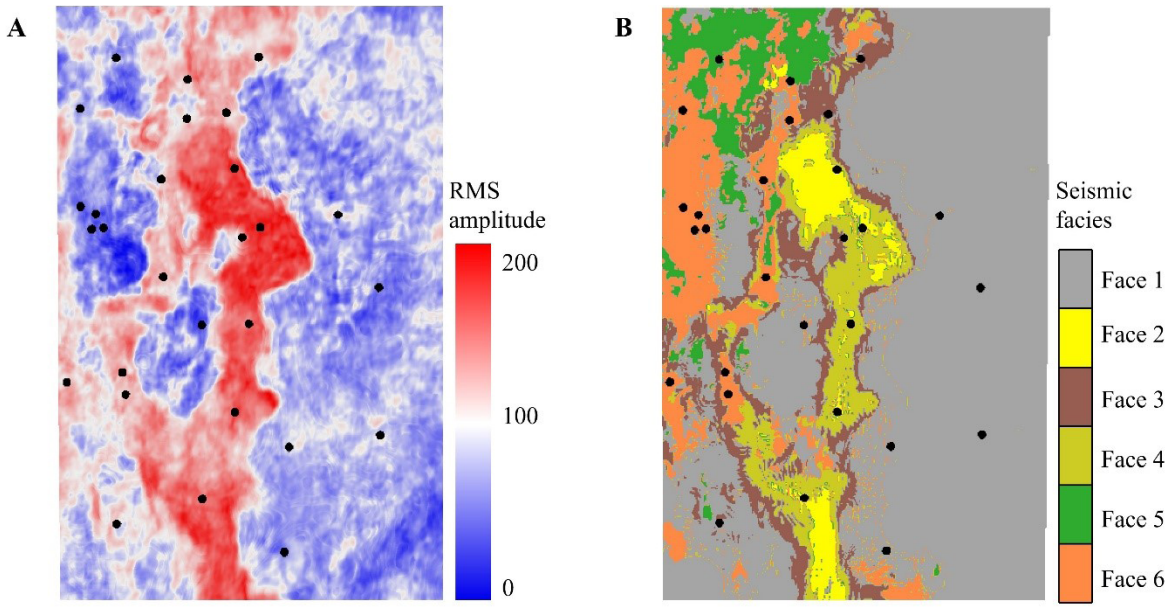



**Figure 8.** Comparison of seismic facies clustering results. (A) Root mean square (RMS) amplitude attribute, showing only two main channels in the central area, with no discernible internal heterogeneity. (B) Seismic facies clustering result, clearly delineating three channels and revealing internal variations. Black dots indicate well locations.

geological bodies typically exhibit strong velocity contrasts and distinct geometric configurations compared to surrounding rocks,[36] producing diverse seismic reflection patterns, including discontinuities, chaotic reflections, and amplitude variations. However, in practice, noise contamination and resolution constraints often obscure geological boundaries, complicating accurate identification. While conventional seismic attributes such as coherence, variance, and fault likelihood have been widely used for detection, they remain sensitive to non-geological discontinuities, frequently generating "fly points" in the identified boundaries and still relying on manual correction due to limited automation. Therefore, accurately characterizing the 3D spatial morphology of specific geological bodies under the constraint of limited manual labels is critical.

To address the scarcity of labeled data, a DL-based image segmentation approach using the U-Net[37-40] architecture was introduced, which treats geological body identification as a semantic segmentation task. In delineating the distribution range of the Zijinshan igneous intrusive body, seismic facies markers for igneous rock identification were first established. These markers are primarily characterized by mound-shaped, moderate to low-amplitude reflections, with localized areas exhibiting oblique reflections relative to the surrounding rocks' seismic phases and strata thickening (Figure 9). Based on these criteria, manual interpretations along only 12 main survey lines, representing approximately 7% of the total dataset, were used as training samples. These interpreted seismic profiles—where igneous rock bodies are assigned a value of 1 and non-igneous rock bodies a value of 0—paired with corresponding raw seismic profiles, served as input–output pairs to train the network. After that, a probability volume of the igneous rock body was predicted for each uninterpreted seismic profile, and the distribution range of the igneous rock body was delineated based on the probability volume. Notably, in contrast to conventional seismic attribute-based methods, which are often limited to two-dimensional planar analysis or qualitative interpretation, our deep learning approach enables the delineation of the 3D spatial extent of such complex geological bodies, thereby providing a spatially continuous and quantitative delineation of the igneous rock distribution with enhanced capability.

The trained network was applied to the entire seismic volume to generate a 3D probability model of the igneous intrusive body in the Zijinshan area. As illustrated in Figure 10, the resulting 3D carving clearly reveals two phases of igneous intrusion. According to the results from *in situ* zircon U–Pb, hornblende and biotite $^{40}Ar$–$^{39}Ar$, and zircon and apatite fission track analyses, Zijinshan intrusive rocks primarily formed during the Late Jurassic to Early Cretaceous, when medium-deep to medium-shallow magmatic-thermal events occurred in the eastern Ordos Basin.[41-43] Within structurally weak zones of the

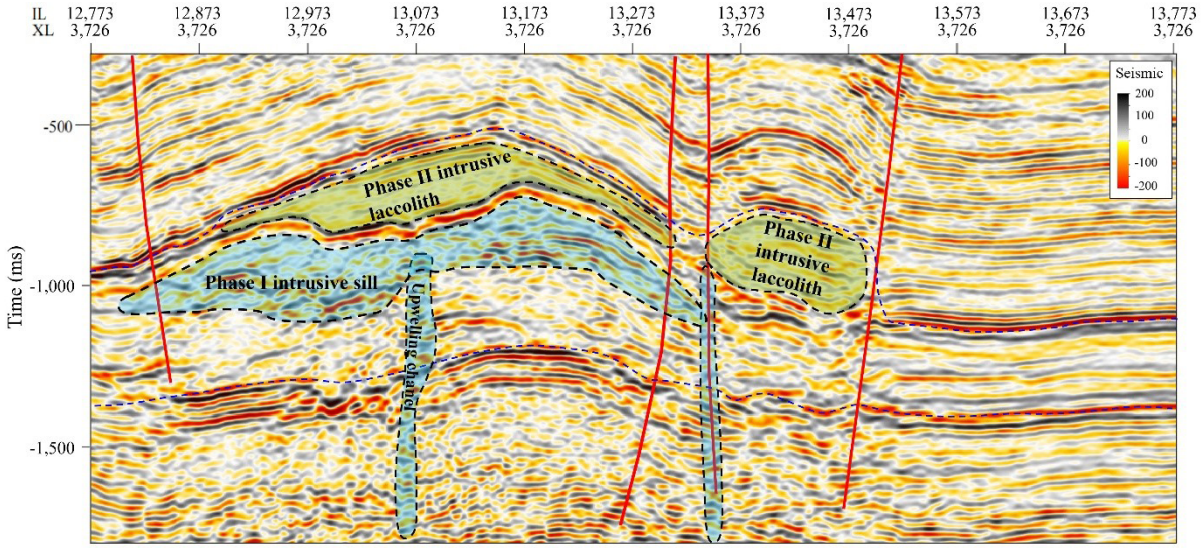



**Figure 9.** Manual interpretation labels for the igneous intrusive body. The red lines denote faults, the blue dashed lines represent horizons, the blue vertical features correspond to magma upwelling channels, the blue irregular shape indicates the Phase I intrusive sill, and the green irregular shape marks the Phase II intrusive laccolith.

Zijinshan area, two distinct magmatic-thermal upwelling events took place, ultimately forming an alkaline complex through mantle-derived intrusion and volcanic eruptions. Furthermore, our results successfully identify these two intrusive cycles (Figure 10).

The first phase consists of sill intrusions primarily emplaced within Ordovician strata, forming the main body of the Zijinshan igneous complex. The second phase forms two asymmetric laccoliths intruding into both Ordovician and Carboniferous layers. Laccolith 1 resulted from magma ascending along faults within the main sill body, while Laccolith 2 formed through magma intrusion along peripheral faults at the edge of the sill. From the perspective of the relationship between the thermal evolution history of the Ordos Basin and hydrocarbon accumulation, the thermal effects associated with magmatism played a crucial

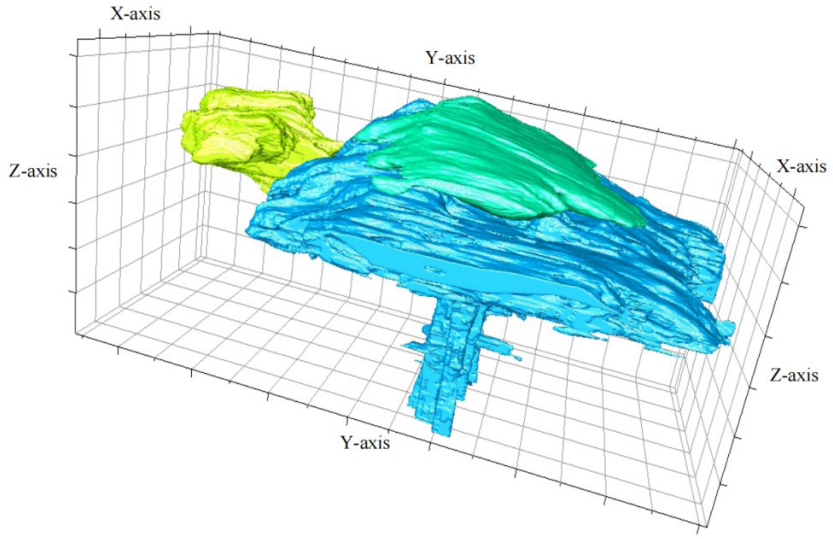



**Figure 10.** Three-dimensional carving results of the igneous intrusive bodies based on semantic segmentation. Blue vertical features correspond to magma upwelling channels, the blue irregular shape indicates the Phase I intrusive sill, and the green and yellow irregular shapes mark the Phase II intrusive laccolith.

role in hydrocarbon expulsion from source rocks.[44]

Additionally, the Zijinshan intrusive contact metamorphic system and its internal configuration show potential for hydrocarbon accumulation. Moreover, due to igneous intrusion and eastern marginal uplift of the basin, the Zijinshan structure zone exhibited exceptionally developed fault systems, providing effective pathways for vertical hydrocarbon migration. Therefore, delineating favorable exploration areas based on the distribution of laccoliths and sills, along with associated fault development patterns, is of significant importance for assessing the multi-gas accumulation potential in the Zijinshan area. The results confirm that the proposed method significantly enhances automated interpretation and accuracy, even with very limited training data, offering an efficient tool for characterizing special geological bodies in complex settings.

### 3.3. Data-driven prediction of reservoir properties

#### *3.3.1. A Bayesian lithofacies classification framework for thin interbeds*

The Linxing–Shenfu Block contains a vertically stacked sequence of marine-continental transitional coal-bearing strata, with main lithologies such as mudstone, tight sandstone, and coal. According to well logging interpretation statistics, the thinnest coal seam measures only 3 m, whereas the thickness of individual sandstone layers in the Benxi and Taiyuan Formations ranges from 2–10 m. The thickness of thin layers is far below the seismic resolution limit, thus posing a challenge to the detailed characterization of reservoirs.

In addition, coal seams exhibit low density and low P-wave impedance, resulting in strong impedance contrasts and high reflection coefficients with surrounding rocks. At low resolution, the strong seismic reflections from coal seams interfere with reflections from adjacent sandstones, resulting in strong amplitude on seismic sections and shielding sandstone reflections near the coal seams, thereby affecting seismic lithofacies identification. Therefore, reliably identifying thin sandstone layers under these constraints represents a key challenge.

Reservoir lithofacies prediction is essential in hydrocarbon reservoir characterization. Current methods commonly use elastic attributes derived from seismic inversion,[45] with techniques including discriminant analysis, Bayesian classification, neural networks, support vector machines, and K-nearest neighbor classification. Among these, Bayesian classification represents the most mature approach, integrating a seismic likelihood function with prior information from well logs to generate posterior probabilities that quantify lithological classification.[46-50] The Bayesian framework also commonly incorporates a one-dimensional Markov chain prior model, formulated as a transition probability matrix, to enforce vertical continuity in lithofacies distribution.[51,52]

Integrating geological constraints with Bayesian methodology, we developed a systematic workflow for lithofacies classification in coal-bearing strata, which mainly includes four steps:

(a) Step 1: Data conditioning and upscaling

Prior to modeling, well logs underwent rigorous quality control, including depth matching (aligning logs with seismic time using checkshot data) and outlier removal to eliminate spikes. To ensure consistency across the survey, all log curves were normalized using *Z*-score scaling. Subsequently, log curve upscaling was performed using the Backus averaging criterion to achieve precise matching between the well and seismic scales.

(b) Step 2: Optimal attribute pair selection

To enhance seismic lithology identification performance in the aliasing region of petro-physical crossplots, a Bayesian classification confusion matrix was introduced to quantify the classification ability of different attribute pairs.

(c) Step 3: High-resolution seismic inversion

To reduce the effects of strong coal seam reflections and identify thin layers, the seismic waveform indicator inversion method was used to obtain high-resolution P-wave impedance and gamma-ray data volumes. The seismic waveform indicator inversion can construct a mapping relationship between seismic traces and high-frequency components of the logging curve, widen the seismic frequency band, and considerably improve the vertical resolution.[53] It also uses horizontal changes in seismic waveforms to reflect lithological assemblage characteristics for facies-controlled constraints. The use of seismic waveforms in inversion further exploits the potential waveform characteristics and wave-group relationships of seismic data, thereby reducing the impact of strong coal-seam reflections.

(d) Step 4: Seismic attribute interpretation

To improve the continuity and accuracy of the interpreted results, a Markov chain-constrained Bayesian classification of seismic attribute volumes was performed, and probability density functions were calculated from well logging curves in step 2 using kernel density estimation, which served as likelihood functions for Bayesian classification.[54] To address the inherent class imbalance—where mudstone is the dominant background and coal/sandstone are minority classes—the Monte Carlo simulation was employed to equally expand the training samples of different lithofacies,

ensuring that each facies has the same sample size in the estimation of conditional probability density functions and in the classification evaluation.

The application in the Linxin–Shenfu Block demonstrates the method's ability to generate detailed 3D lithofacies predictions. Figure 11 shows that our approach successfully identified not only coal seams but also thin sandstone layers immediately adjacent to them. Statistically, the lithology predictions obtained by this method showed high consistency with lithology interpretations derived from well logs. The method can identify coal seams as thin as 3.3 m, and the recognition accuracy for coal seams, tight sandstone, and mudstone is 0.90, 0.86, and 0.85, respectively. The derived lithofacies volume enables the extraction of stratigraphic attributes for mapping coal distribution and characterizing sandstone reservoirs, thereby providing substantial technical support for predictive reservoir modeling.

### *3.3.2. Integrating geological priors in deep learning for physical property prediction*

For tight sandstone and shale gas reservoirs characterized by low porosity, low permeability, and strong heterogeneity, identifying localized areas with high porosity, high permeability, and high gas saturation, commonly referred to as "sweet spots," is crucial for enhancing hydrocarbon productivity.[55] Accurate delineation of these "sweet spots" can avoid inefficient well placements and significantly improve development economics. Extensive drilling and gas test results in the Linxing–Shenfu Block indicate that high-productivity wells are typically associated with reservoirs with porosity greater than 12%, permeability exceeding 1 mD, and gas saturation above 50%. Consequently, the quantitative prediction of these "sweet spots" is our primary objective.

Conventional methods for predicting physical properties primarily include attribute analysis,[56,57] frequency-division

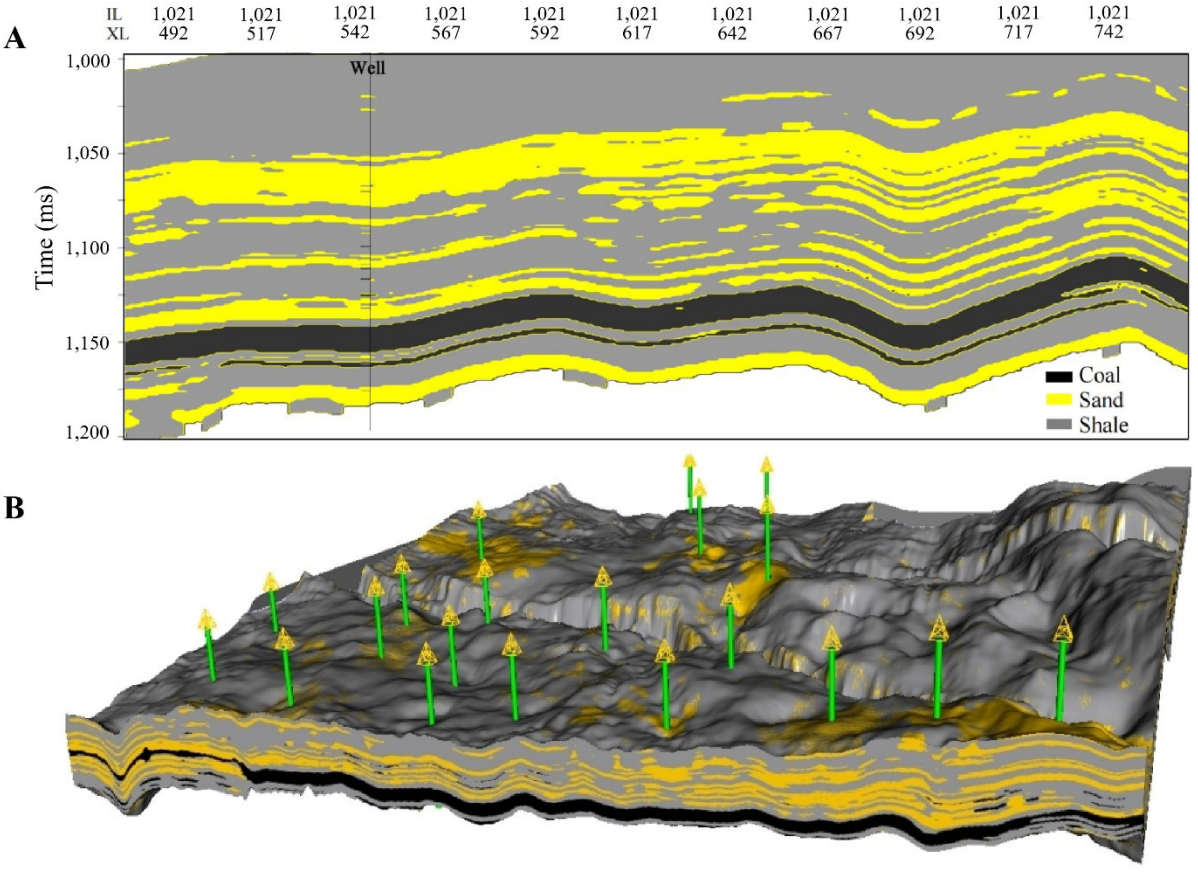



**Figure 11.** Lithology results based on Bayesian classification prediction, with yellow, black, and gray representing sandstone, coal, and mudstone, respectively. (A) Lithology profiles. (B) Three-dimensional lithology volume, with drilled wells indicated in green.

attribute fusion,[58] and seismic inversion.[59,60] For instance, by analyzing differences in petrophysical properties or seismic response characteristics between sweet-spot and non-sweet-spot reservoirs, inversion techniques and attribute analysis were employed to identify sweet-spot distribution. However, these traditional techniques face dual challenges. On the one hand, these methods largely rely on empirical relationships or qualitative analysis of regional data. Due to the lack of a direct analytical relationship between reservoir physical parameters and seismic data, it is difficult to establish a direct mapping between them.[61] On the other hand, our study block exhibited a significant contradictory sample phenomenon (i.e., similar seismic waveforms correspond to drastically different well log responses, or different reservoir parameters correspond to similar seismic response characteristics). This non-unique correspondence between seismic and well log data severely restricts the direct application of conventional methods, leading to unstable prediction results, insufficient vertical resolution, and an inability to meet the requirements for finely characterizing sweet-spot distribution during the development stage.[62]

In addressing such geological challenges, many researchers have turned to DL techniques—training networks to learn the implicit relationships that map seismic data to physical property parameters.[63] By constructing complex network architectures, such as CNNs, recurrent neural networks, deep belief networks, and hybrid DL networks, the complex nonlinear relationships between seismic data and reservoir parameters can be explored. This approach fully leverages the advantages of a data-driven methodology, bypasses the seismic rock physics modeling step, and improves reservoir prediction accuracy. Promising results for predicting reservoir physical properties have been achieved in areas with high well density. Building upon this, our study introduces a further improvement by innovatively proposing a DL-based technical framework integrated with geological prior constraints.[62] Specifically, our DL-based network model enhances the conventional CNN by incorporating fully connected layers. These layers integrate prior information, such as horizons and seismic facies, to resolve contradictory samples and ensure stable predictions. Simultaneously, the raw seismic data were fed into a locally connected CNN. The outputs from these two parts were then fused to produce the network's final output. The training labels (e.g., porosity, permeability, and gas saturation) were derived from log interpretations rigorously calibrated against core analysis data.

Accordingly, we constructed a training dataset using seismic data and geological prior information as inputs and the reservoir parameters (e.g., porosity, permeability, and gas saturation) from sample wells as outputs to train the network. During the screening of sample wells, we established strict geologically guided criteria, ensuring not only uniform spatial coverage across the study area but also representation of key sedimentary microfacies, including channel bars, composite channels, channel margins, and inter-channel bays, thereby guaranteeing sample balance and diversity. Application of this network in the Linxing–Shenfu Block demonstrates that the predicted physical properties not only match well with drilling data but also exhibit stronger geological consistency (Figure 12), thereby effectively delineating the distribution of "sweet spots."

#### *3.3.3. Gas-bearing property discrimination using a multi-attribute neural network*

Gas-bearing property prediction, the ultimate objective of seismic reservoir characterization, constitutes a highly non-unique problem due to limitations in seismic SNR, resolution, and complex rock physics. Conventional approaches relying on seismic attribute analysis (e.g., amplitude, frequency,[64] attenuation,[65] and amplitude-versus-offset) or seismic inversion techniques[66] exhibit limited discrimination capability for thin interbedded layers, low-SNR data, or complex lithological structures. Moreover, the complex dependence of gas-bearing properties on multiple factors, including rock physics, stratigraphic architecture, and fluid saturation, cannot be adequately captured by individual attributes or conventional analytical methods.

The eastern part of the Linxing–Shenfu Block, tectonic uplift associated with the Luliang Uplift, has led to diminished hydrocarbon generation intensity in source rocks and insufficient gas charging dynamics, resulting in widespread gas-water mixing within the Shihezi Formation. Post-stack seismic profiles showed no significant differences between gas-bearing and water-bearing layers, while rock physics analysis revealed overlapping elastic parameters. These conditions render conventional discrimination methods ineffective, making reliable gas prediction a critical challenge in this area.

The TG curve—one of the primary logging methods—contains substantial reservoir information through its amplitude values and morphological variations, serving as a direct indicator for detecting and evaluating gas-bearing zones.[67] Gas-bearing layers typically exhibit high TG values, whereas water-bearing layers show low values. While providing high vertical resolution, TG lacks lateral continuity, and seismic data exhibit opposite character. By integrating the complementary advantages of these two data types through an appropriate methodology, direct prediction of reservoir gas-bearing potential can

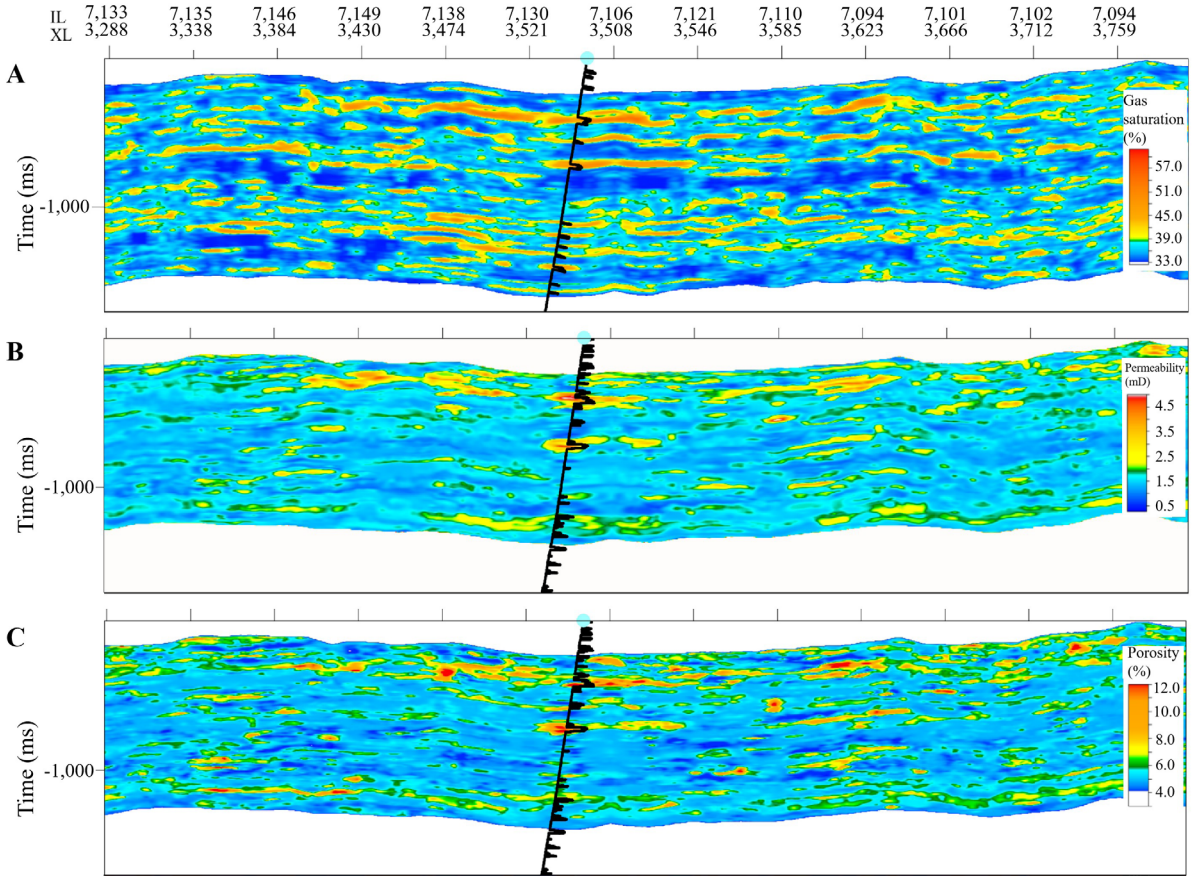



**Figure 12.** Prediction results of physical properties: (A) Gas saturation, (B) permeability, and (C) porosity

be achieved. DL technology enables effective integration of these complementary datasets and geological pattern recognition,[68,69] establishing nonlinear mappings between multi-attribute seismic inputs and TG responses to resolve prediction non-uniqueness.

To address these challenges, a workflow with a multi-layer feed-forward neural network was employed. The workflow initiates with standardized correction of TG from well logs, accounting for influencing factors such as drill bit diameter, drilling time, and drilling fluid density. This correction ensures that the normalized TG values reliably reflect vertical variations in gas-bearing properties while maintaining lateral comparability across the study area. Subsequently, the Hilbert transform was applied to the seismic data to extract a comprehensive set of attributes, including amplitude, frequency, and phase components within the target interval, generating a seismic attribute database. Pearson correlation coefficients were then calculated between the corrected TG and each seismic attribute at well locations. Attributes showing significant correlation with TG were retained, while those inconsistent with geological understanding were systematically eliminated. The optimized attribute set, together with the corrected TG data, served as input to a neural network. Model training was performed under the constraint of pre-stack P- and S-wave velocity ratio (Vp/Vs) inversion results to reinforce petrophysical consistency, thereby establishing a robust nonlinear mapping relationship between the multi-attribute seismic input and TG properties. This integrated approach enhances prediction reliability by jointly leveraging the complementary strengths of well logs, seismic attributes, and rock physics constraints.

The trained model was applied to the full 3D seismic volume to generate a continuous TG prediction cube. This resultant data cube enables delineation of gas-bearing distributions within the target formation, informed by structural interpretation. Notably, the TG prediction successfully identified gas zones that are indistinguishable from water-bearing layers using conventional pre-stack inversion methods, demonstrating effective gas-layer prediction even under high water-saturation conditions (Figure 13).

## 4. Key challenges and future directions in unconventional reservoir characterization

Based on the practices in the Linxing–Shenfu Block, several key research directions for intelligent geophysics are summarized as follows:

(i) Emphasize the core role of data quantity and quality in improving ML models: The Linxing–Shenfu Block remains in the early stage of oil and gas exploration and development, and data are still being accumulated. Through continuous data accumulation and cleaning, noise removal, and SNR improvement, the models are enabled to learn more representative geological features from purer and richer data. Moreover, iterative data utilization, that is, fine-tuning models based on feedback from prediction results after initial model training and continuous model iteration, further optimizes prediction accuracy, thereby highlighting the critical role of data quantity and quality in model optimization.

(ii) ML models should avoid excessive complexity: Compared to data from the complex physical world, geoscientific data exhibit cyclicity and low dimensionality. The design of intelligent geophysical models follows the Occam's Razor principle, that is, models should be as simple and interpretable as possible while meeting interpretation requirements. From current practice, various improved complex intelligent models provide limited improvements in efficiency and accuracy for industrial-level

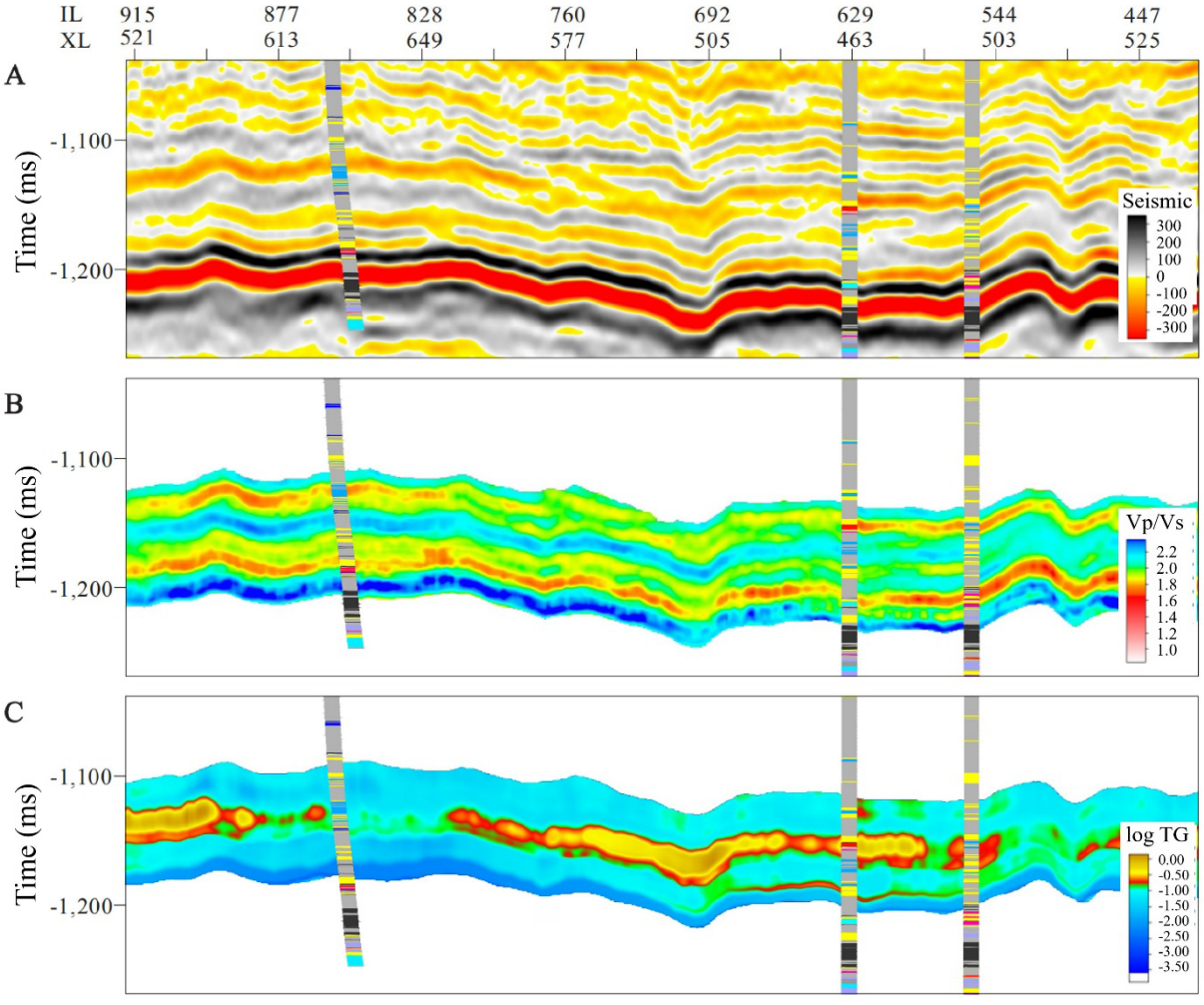



**Figure 13.** Prediction results of total gas (TG). (A) Seismic data. (B) Pre-stack P- and S-wave velocity ratio (Vp/Vs) obtained from pre-stack inversion. (C) TG prediction results, with red, blue, yellow, and gray at well locations representing gas layers, water layers, sandstone, and mudstone, respectively.

interpretation tasks, but pose high requirements on data input and model iterative maintenance.

(iii) The integration of prior geological and geophysical knowledge with ML models: Fundamental geological and geophysical understanding exerts a strong bias effect on ML models. In intelligent geophysical research, it is essential to fully integrate geologists' prior knowledge and domain expertise to ensure that models better simulate and predict real geological phenomena while maintaining a basic generalization ability. By incorporating geological logic into algorithm design, such as combining CRFs and path-optimization algorithms in horizon interpretation, both geological laws and the automation advantages of DL are considered. During model training, expert-curated, geologically significant labeled data are introduced to align predictions with geological reality, thereby enabling seamless integration of geological knowledge with intelligent algorithms.

## 5. Conclusion

This study presented a comprehensive, ML-driven workflow for seismic interpretation, applied and validated in the complex Linxing–Shenfu tight gas and CBM field. The key achievements are summarized as follows:

(i) An integrated intelligent workflow: We established and implemented a seamless, data-driven interpretation chain. This workflow systematically progresses from structural framework building through intelligent horizon and fault interpretation under low-SNR conditions to stratigraphic analysis via deep clustering for seismic facies and semantic segmentation of special geological bodies (e.g., the Zijinshan igneous rock mass).

(ii) Quantitative reservoir characterization: Leveraging deep learning models constrained by geological priors, this workflow successfully transitions from qualitative interpretation to quantitative prediction. It enables the direct and accurate estimation of key reservoir properties, including lithology, porosity, permeability, and gas saturation, culminating in the fine-scale, 3D characterization of composite geological "sweet spots."

(iii) Practical efficacy and broader implications: The application of this workflow has not only significantly enhanced interpretation efficiency and objectivity but has also effectively addressed persistent theoretical bottlenecks of conventional methods. These include managing low-SNR data, identifying thin interbeds, and resolving fluid ambiguity in complex rock physics scenarios. The success in the Linxing–Shenfu Block underscores the paradigm-shifting potential of integrated artificial intelligence-based workflows for unlocking the value of complex unconventional reservoirs, offering a robust and scalable template for efficient exploration and development decision-making in geologically analogous plays worldwide.

## Acknowledgments

None.

## Funding

None.

## Conflict of interest

The authors declare no potential conflict of interest.

## Author contributions

*Conceptualization:* Qixin Li
*Formal analysis:* Xiaowen Zheng, Bo Wang
*Investigation:* Di Wang, Jicai Ding
*Methodology:* Qixin Li
*Writing–original draft:* Wenlan Li
*Writing–review & editing:* Wenlan Li, Bo Wang

## Availability of data

Data are available from the corresponding author upon reasonable request.

## References

1. Liu JZ, Zhu GH, Liu YC, *et al*. Breakthrough, future challenges and countermeasures of deep coalbed methane in the eastern margin of Ordos Basin: A case study of Linxing-Shenfu block. *Acta Pet Sin*. 2023;44(11):1827-1839. [In Chinese].

   doi: 10.7623/syxb202311006

2. Mi L, Zhu G. Geological characteristics and exploration breakthrough in Linxing-Shenfu tight gas field, northeastern Ordos Basin. *China Pet Explor*. 2021;26(3):53-67. [In Chinese].

   doi: 10.3969/j.issn.1672-7703.2021.03.005

3. Wang SX, Yuan SY. *Ren gong zhi neng di qiu wu li kan tan. Artificial Intelligence for Geophysical Exploration*. Beijing: Science Press; 2025. [In Chinese].

4. Yu SW, Ma JW. Deep learning for geophysics: Current and future trends. *Rev Geophys*. 2021;59(3):e2021RG000742.

   doi: 10.1029/2021RG000742

5. Zhu YH, Zhao ZG, Lu YY, *et al*. Research progress and key technologies for exploration of onshore tight sandstone gas by CNOOC. *China Offshore Oil Gas*. 2024;36(6):13-25. [In Chinese].

doi: 10.11935/j.issn.1673-1506.2024.00.007

6. Zhu GH, Li BL, Li ZC, *et al*. Practices and development trend of unconventional natural gas exploration in eastern margin of Ordos Basin: Taking Linxing-Shenfu gas field as an example. *China Offshore Oil Gas*. 2022;34(4):16-29. [In Chinese].

   doi: 10.11935/j.issn.1673-1506.2022.04.002

7. Liu C, Zhang DM, Li C, Lu YY, Guo MQ. Upper Paleozoic tight gas sandstone reservoirs and main controls, Linxing block, Ordos Basin. *Oil Gas Geol*. 2021;42(5):1146-1158. [In Chinese].

   doi: 10.11743/ogg20210512

8. Ge Y, Zhu GH, Wan H, Pan XZ, Huang ZL. The influence of Zijinshan Structural belt to the formation and distribution of tight sandstone gas reservoir in Upper Paleozoic, in the eastern Ordos Basin. *Nat Gas Geosci*. 2018;29(4):491-499. [In Chinese].

   doi: 10.11764/j.issn.1672-1926.2018.03.008

9. Bergen KJ, Johnson PA, de Hoop MV, Beroza GC. Machine learning for data-driven discovery in solid Earth geoscience. *Science*. 2019;363(6433):eaau0323.

   doi: 10.1126/science.aau0323

10. Herron DA. Pitfalls in horizon autopicking. *Interpretation*. 2015;3(1):SB1-SB4.

    doi: 10.1190/INT-2014-0062.1

11. Hoyes J, Cheret T. A review of 'global' interpretation methods for automated 3D horizon picking. *Lead Edge*. 2011;30(1):38-47.

    doi: 10.1190/1.3535431

12. Wu XM, Fomel S. Least-squares horizons with local slopes and multi-grid correlations. *Geophysics*. 2018;83(4):IM29-IM40.

    doi: 10.1190/geo2017-0830.1

13. Xu YM. Di zhen pou mian duan ceng de jie shi yan jiu. Interpretation of faults in seismic profiles. *China Pet Chem Stand Qual*. 2012;33(10):50. [In Chinese].

    doi: 10.3969/j.issn.1673-4076.2012.10.040

14. Xi GM, He SG, Min Y, Sun YC, Li QC. Fault identification by the coherency attributes. *Northwest Geol*. 2019;52(1):244-249. [In Chinese].

    doi: 10.19751/j.cnki.61-1149/p.2019.01.023

15. Cai HP. Improvement of variance cube technique and its application in seismic interpretation. *Coal Geol Explor*. 2008;36(1):74-76. [In Chinese].

    doi: 10.3969/j.issn.1001-1986.2008.01.019

16. He Y, He ZH, Xiong XJ. Fault Identification Based on High-precision Curvature Analysis. *J Oil Gas Technol*. 2010;32(6):404-407. [In Chinese].

    doi: CNKI:SUN:JHSX.0.2010-06-102

17. Li HW, Bai XL, Cui JB, Wan HZ, Yuan SH, Chu WC. Fault identification technology of ant attribute optimization. *Coal Geol Explor*. 2019;47(6):174-179. [In Chinese].

    doi: 10.3969/j.issn.1001-1986.2019.06.026

18. Bi ZF, Wu XM, Geng ZC, Li HS. Deep relative geologic time: A deep learning method for simultaneously interpreting 3-D seismic horizons and faults. *J Geophys Res Solid Earth*. 2021;126(9):e2021JB021882.

    doi: 10.1029/2021JB021882

19. Geng ZC, Wu XM, Shi YZ, Fomel S. Deep learning for relative geologic time and seismic horizons. *Geophysics*. 2020;85(4):WA87-WA100.

    doi: 10.1190/geo2019-0252.1

20. Luo YL, Zhang GL, Li L, Zhang XD, Duan J, Li XW. Attention-based two-stage U-Net horizon tracking. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

    doi: 10.1109/LGRS.2022.3217376

21. Wu XM, Liang LM, Shi YZ, Fomel S. FaultSeg 3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophysics*. 2019;84(3):IM35-IM45.

    doi: 10.1190/geo2018-0646.1

22. Li Y, Wu XM, Zhu ZY, Ding JC, Wang QZ. FaultSeg 3D plus: a comprehensive study on evaluating and improving CNN-based seismic fault segmentation. *Geophysics*. 2024;89(5):N77-N91.

    doi: 10.1190/geo2022-0778.1

23. Wei XL, Zhang CX, Kim SW, *et al*. Seismic fault detection using convolutional neural networks with focal loss. *Comput Geosci*. 2022;158:104968.

    doi: 10.1016/j.cageo.2021.104968

24. Li SZ, Liu NH, Li FY, Gao JH, Ding JC. Automatic fault delineation in 3-D seismic images with deep learning: Data augmentation or ensemble learning? *IEEE Trans Geosci Remote Sens*. 2022;60:1-14.

    doi: 10.1109/TGRS.2022.3150353

25. Zhou ZW, Rahman Siddiquee MM, Tajbakhsh N, Liang JM. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *arXiv*. Preprint posted online July 18, 2018. Accessed March 19, 2021.

    doi: 10.48550/arXiv.1807.10165

26. Liu BG, Liu LH. Application of Applied Seismic Sedimentology in Sedimentary Facies Analysis. *Geophys Prospect Pet*. 2008;47(3):266-271. [In Chinese].

    doi: 10.3969/j.issn.1000-1441.2008.03.010

27. Wrona T, Pan I, Gawthorpe RL, Fossen H. Seismic facies analysis using machine learning. *Geophysics*.

2018;83(5):O83-O95.

doi: 10.1190/geo2017-0595.1

28. Zhang LK, Qin LJ, Zhang GS, Sun XQ. On Seismic Facies Analysis Based on Seismic Attributes. *Chin J Eng Geophys*. 2010;7(6):694-698. [In Chinese].

doi: 10.3969/j.issn.1672-7940.2010.06.009

29. Barnes A. Redundant and useless seismic attributes. *Geophysics*. 2007;72(3):P33-P38.

doi: 10.1190/1.2716717

30. Yan XY, Gu HM, Luo HM, Yan YP. Intelligent Seismic facies classification based on an improved deep learning method. *Oil Geophys Prospect*. 2020;55(6):1169-1177. [In Chinese].

doi: 10.13810/j.cnki.issn.1000-7210.2020.06.001

31. Shuo LX, Zhao YH, Chai BF, Li ZH, Wu H. Semi-supervised adversarial network for seismic facies classification. *Prog Geophys*. 2023;38(5):2105-2113. [In Chinese].

doi: 10.6038/pg2023GG0372

32. Qian F, Yin M, Liu XY, Wang YJ, Lu C, Hu GM. Unsupervised seismic facies analysis via deep convolutional autoencoders. *Geophysics*. 2018;83(3):A39-A43.

doi: 10.1190/geo2017-0524.1

33. Liu ML, Jervis M, Li WC, Nivlet P. Seismic facies classification using supervised convolutional neural networks and semisupervised generative adversarial networks. *Geophysics*. 2020;85(4):O47-O58.

doi: 10.1190/geo2019-0627.1

34. Li QX, Luo YN, Ma XQ, Chen C, Zhu YH. Seismic facies clustering technology based on deep embedding network. *Oil Geophys Prospect*. 2022;57(2):251-260. [In Chinese].

doi: 10.13810/j.cnki.issn.1000-7210.2022.02.002

35. Wu XM, Yang JR, Zhu ZY, Ding JC, Wang QZ. Discussions on computational seismic structural interpretation and modeling. *Geophys Prospect Pet*. 2022;61(3):392-407. [In Chinese].

doi: 10.3969/j.issn.1000-1441.2022.03.002

36. Jones IF, Davison I. Seismic imaging in and around salt bodies. *Interpretation*. 2014;2(4):SL1-SL20.

doi: 10.1190/INT-2014-0033.1

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

doi: 10.1007/978-3-319-24574-4_28

38. Shi YZ, Wu XM, Fomel S. Automatic salt-body classification using a deep convolutional neural network. In: *SEG Technical Program Expanded Abstracts 2018*. SEG International Exposition and Annual Meeting 2018; August 1, 2018. Society of Exploration Geophysicists; 2018:1971-1975.

doi: 10.1190/segam2018-2997304.1

39. Di HB, Wang Z, AlRegib G. Deep convolutional neural networks for seismic salt-body delineation. In: *Proceedings of the AAPG 2018 Annual Convention & Exhibition*. AAPG 2018 Annual Convention & Exhibition; May 20–23, 2018; Salt Lake City, Utah, United States. AAPG; 2018.

doi: 10.1306/70630Di2018

40. Shi YZ, Wu XM, Fomel S. SaltSeg: Automatic 3D salt segmentation using a deep convolutional neural network. *Interpretation*. 2019;7(3):SE113-SE122.

doi: 10.1190/INT-2018-0235.1

41. Yang XK, Yang YH, Ji LD, Su CQ, Zheng ML, Zhao L. Stages and characteristics of thermal actions in eastern part of Ordos Basin. *Acta Geol Sin*. 2006;80(5):705-711. [In Chinese].

doi: 10.3321/j.issn:0001-5717.2006.05.010

42. Yang XK, Chao HX, Zhang ZF, Yao WH, Dong M. Characteristics of the Zijinshan complex and its dynamic environment in the east of the Ordos Basin—Analysis of a classic example on the deep processes of thermal-magma activity in the basin. *Geotecton Metallog*. 2010;34(2):269-281. [In Chinese].

doi: 10.16539/j.ddgzyckx.2010.02.008

43. Chen G, Ding C, Xu LM, *et al*. Analysis on the thermal history and uplift process of Zijinshan intrusive complex in the eastern Ordos basin. *Chin J Geophys*. 2012;55(1):78-90. [In Chinese].

doi: 10.6038/j.issn.0001-5733.2012.11.020

44. Ren ZL, Zhang S, Gao SL, Cui JP, Xiao YY, Xiao H. E'erduosi pen di gou zao re yan hua shi ji qi cheng kuang cheng cang yi yi. Tectono-thermal evolution history of the Ordos Basin and its implications for hydrocarbon accumulation and mineralization. *Sci China Press*. 2007;37(suppl 1):23-32. [In Chinese].

45. Cooke DA, Schneider WA. Generalized linear inversion of reflection seismic data. *Geophysics*. 1983;48(6):665-676.

doi: 10.1190/1.1441497

46. Mukerji T, Avseth P, Mavko G, Takahashi I, Gonzalez EF. Statistical rock physics: combining rock physics, information theory, and geostatistics to reduce uncertainty in seismic reservoir characterization. *Lead Edge*. 2001;20(3):313-319.

doi: 10.1190/1.1438938

47. Avseth P, Mukerji T, Mavko G. *Quantitative Seismic Interpretation*. Cambridge University Press; 2005.

48. Grana D. Joint facies and reservoir properties inversion. *Geophysics*. 2018;83(3):M15-M24.

doi: 10.1190/GEO2017-0670.1

49. Grana D. Bayesian linearized rock-physics inversion. *Geophysics*. 2016;81(6):D625-D641.

doi: 10.1190/GEO2016-0161.1

50. Aleardi M, Ciabarri F. Application of different classification methods for litho-fluid facies prediction: a case study from the offshore Nile Delta. *J Geophys Eng*. 2017;14(5):1087-1102.

doi: 10.1088/1742-2140/aa7301

51. Larsen AL, Ulvmoean M, Omre H, Buland A. Bayesian lithology/fluid prediction and simulation on the basis of a Markov-Chain prior model. *Geophysics*. 2006;71(5):R69-R78.

doi: 10.1190/1.2245469

52. Wang L, Zhang F, Li XY, Di BR, Zeng LB. Quantitative seismic interpretation of rock brittleness based on statistical rock physics. *Geophysics*. 2019;84(4):IM63-IM75.

doi: 10.1190/geo2018-0094.1

53. Kong W, Zhang LJ, Cheng MW, Zhang HW, Jiang LZ. Application of Waveform Indication Inversion in the Prediction of Thin Carbonate Reservoirs in the Amu Darya Right Bank Area. In: Lin J, ed. *Proceedings of the International Field Exploration and Development Conference 2019*. IFEDC 2019; October 16–18, 2019; Xi'an, China. Springer; 2020:3256-3266.

doi: 10.1007/978-981-15-2485-1_302

54. Qi Y, Wu K, Wang B, Zheng XW, Li WL, Li D. Lithofacies identification of deep coalbed methane reservoir based on high-resolution seismic inversion. *Front Earth Sci*. 2024;12:1440729.

doi: 10.3389/feart.2024.1440729

55. Sena A, Castillo G, Chesser K, *et al*. Seismic reservoir characterization in resource shale plays: 'sweet spot' discrimination and optimization of horizontal well placement. In: *SEG Technical Program Expanded Abstracts 2011*. 81th SEG Annual International Meeting; September 18–23, 2011; San Antonio, Texas, United States. Society of Exploration Geophysicists; 2011:1744-1748.

doi: 10.1190/1.3627542

56. Hart BS, Pearson R, Rawling GC. 3-D seismic horizon-based approaches to fracture-swarm sweet spot definition in tight-gas reservoirs. *Lead Edge*. 2002;21(1):28-35.

doi: 10.1190/1.1445844

57. Le J, Jing YQ, Liang X, Fan HJ, Cai WT. Using seismic response characteristics to predict "sweet spot" locations in a low-permeability reservoir in the Bohai Sea. *Geophys Prospect Pet*. 2022;61(5):876-887. [In Chinese].

doi: 10.3936/j.issn.1000-1441.2022.05.012

58. Liu ZQ, Wang YM, Bai GC, Gong CL. Application of sweetness and its fusion attribute in the research of Deep-water reservoir. *Oil Geophys Prospect*. 2010;45(suppl 1):158-162.

doi: 10.13810/j.cnki.issn.1000-7210.2010.s1.005

59. Peng CZ, Peng J, Chen YH, Zhang HR. Seismic predicting of sweet spots in the Da'anzhai Shale Play, Yuanba area, Sichuan Basin. *Nat Gas Ind*. 2014;34(6):42-47. [In Chinese].

doi: 10.3787/j.issn.1000-0976.2014.06.007

60. Zhao C, Jin FM, Han GM, Guo SW, Xing X, Liu HZ. Direct prediction of sweet spots in sandstone reservoirs based on pre-stack probability inversion. *Oil Geophys Prospect*. 2023;58(5):1211-1219. [In Chinese].

doi: 10.13810/j.cnki.issn.1000-7210.2023.05.018

61. Qian YG. Application of machine deep learning technology in tight sandstones reservoir prediction: A case study of Xujiahe Formation in Xinchang, western Sichuan Depression. *Pet Reserv Eval Dev*. 2023;13(5):600-607. [In Chinese].

doi: 10.13809/j.cnki.cn32-1825/te.2023.05.007

62. Wang D, Zhang YM, Zhang FC, Ding JC, Niu C. Quantitative prediction of tight sandstone sweet spots based on deep learning method with prior information constraints. *Oil Geophys Prospect*. 2023;58(1):65-74. [In Chinese].

doi: 10.13810/j.cnki.issn.1000-7210.2023.01.006

63. An P, Cao DP. Ji yu tan xing zu kang he shen du xue xi de mi du yu ce fang fa yan jiu. Research on density prediction methods based on elastic impedance and deep learning. In: Proceedings of Annual Meeting of Chinese Geoscience Union 2018, October 21–23, 2018; Beijing, China. China Peace Publishing House; 2018:754-757. [In Chinese].

64. Zhang MG, Hong Z, Dou YT, Cui XJ. Application of time-frequency analysis technology to the gas detection in Sulige area. *Lithol Reserv*. 2013;25(5):76-85. [In Chinese].

65. Zhang YQ, Wang ZZ, Zhang N. Research on sand body identification and gas bearing area prediction of He 8 member in SU X Well Block of Sulige Gas Field. *Nat Gas Geosci*. 2011;22(1):164-170. [In Chinese].

66. Lin LM, Zheng Y, Shi H, *et al*. Gas-bearing prediction of thin tight sandstone reservoirs based on facies-controlled geostatistical pre-stack inversion: A case study of the middle Linxing Block in the Ordos Basin. *Nat Gas Ind*. 2023;43(2):56-66. [In Chinese].

doi: 10.3787/j.issn.1000-0976.2023.02.006

67. Deng BQ, Ding D. Reservoir prediction method based on seismic-gas logging fusion. *Mud Logging Eng*. 2025;36(1):77-84. [In Chinese].

doi: 10.3969/j.issn.1672-9803.2025.01.012

68. Song ZH, Yuan SY, Li ZM, Wang SX. KNN-based gas-

bearing prediction using local waveform similarity gas-indication attribute - An application to a tight sandstone reservoir. *Interpretation*. 2022;10(1):T1-T10.

doi: 10.1190/INT-2021-0045.1

69. Gao JH, Song ZH, Gui JY, Yuan SY. Gas-bearing prediction using transfer learning and CNNs: An application to a deep tight dolomite reservoir. *IEEE Geosci Remote Sens Lett*. 2022;19:3001005.

doi: 10.1109/LGRS.2020.3035568

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# XGBoost-Deep Residual (FedXGB-ResNet) collaborative porosity prediction in a federated learning framework

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

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

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

[5]National Observation and Research Station for Wuhan Gravitation and Solid Earth Tides, Hubei Earthquake Administration, Wuhan, Hubei, China

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

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

***Corresponding author:**
Piyapatr Busababodhin
(piyapatr.b@msu.ac.th)

**Citation:** Zhao T, Chen G, Li J, *et al.* XGBoost-Deep Residual (FedXGB-ResNet) collaborative porosity prediction in a federated learning framework. *J Seismic Explor.* 2026;35(2):025440094. doi: 10.36922/JSE025440094

**Received:** October 27, 2025

**Revised:** January 9, 2026

**Accepted:** January 13, 2026

**Published online:** April 16, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Abstract

Privacy protection and multimodal fusion represent significant challenges in the context of cross-institutional geological data collaboration. To address these challenges, this paper proposes a collaborative XGBoost-Deep Residual model (FedXGB-ResNet) within a federated learning framework to achieve high-accuracy porosity prediction. Through a heterogeneous federated integration architecture, the framework combines XGBoost's efficient modeling of structured well-logging parameters (with a feature importance gain of 41.2%) with ResNet's deep extraction of spatiotemporal features from seismic images (87% overlap in activation maps between training and validation sets). The model achieved R2 scores of 0.87 and 0.83 on North Sea and Bakken oilfield datasets, respectively, representing improvements of 12.7% and 9.3% over the traditional FedAvg-XGBoost baseline. The innovative dual privacy protection mechanisms—gradient obfuscation and Paillier encryption—suppressed the membership inference attack success rate to 13.7% and reduced gradient similarity to 0.19 at a noise scale of 1.5 (privacy budget = 0.75), while increasing communication time by only 23%. A dynamic feature distillation mechanism adaptively fuses multimodal features through gated attention units, narrowing the F1-score gap between high-porosity and low-porosity identification to 5%. Experiments demonstrated that the framework reduces the risk of privacy leakage by 84.7% while retaining 93.4% of the performance of a centralized model, offering a balanced solution for collaborative cross-domain geological data analysis in terms of accuracy, privacy, and efficiency.

***Keywords:*** Federated learning; Porosity prediction; XGBoost; Deep residual network; Multimodal fusion; Geological data analysis

## 1. Introduction

### 1.1. Background and motivation

In the current context of digital energy exploration, cross-institutional geological data collaboration faces severe technical and compliance challenges. Increasingly stringent global data privacy regulations (such as the "right to be forgotten" in Article 17 of the General Data Protection Regulation [GDPR] and Section 1798.150 of the California Consumer Privacy Act [CCPA]) have called into question the legitimacy of traditional centralized modeling approaches. The inherent data silos in geological exploration further exacerbate this conflict. Significant distribution shifts in well-logging parameters exist, such as a 25.4% difference in mean resistivity (RT) between North American shale gas fields and Middle Eastern carbonate oil fields. However, commercial competition has led to a lack of data-sharing mechanisms. Existing porosity prediction research is largely limited to data from a single institution. While centralized models based on convolutional networks have been used for analysis in the Bakken oilfield, their data integration methods directly violate the cross-border data transfer provisions of Section 5 of the EU Data Governance Directive. This dilemma highlights the urgent need for a new paradigm in geological data analysis.

### 1.2. Literature review

The application of federated learning (FL) across various domains demonstrates its potential for privacy-preserving collaborative modeling. From a rock physics perspective, porosity prediction requires integrating well-established petrophysical relationships[1-3] with modern machine learning approaches to ensure geological consistency. Lin *et al.*[4] introduced a federated matrix factorization framework, MetaMF, aimed at rating prediction in mobile environments, highlighting the adaptability of FL to different data modalities. Similarly, Chen *et al.*[5] verified the feasibility of horizontal FL for the quantitative structure–activity relationship analysis in drug discovery, emphasizing the framework's capacity for collaborative chemical property prediction without data sharing.

In the realm of location data, Yin *et al.*[6] reviewed FL algorithms and models, such as deep neural networks and Gaussian processes, underscoring the versatility of FL in data-driven cooperative localization. Lu *et al.*[7] extended FL to computational pathology, employing weakly supervised attention multiple instance learning and differential privacy to analyze gigapixel whole slide images, demonstrating FL's effectiveness in high-dimensional medical imaging.

Addressing challenges related to model calibration and interpretability, Lu and Kalpathy-Cramer[8] proposed integrating conformal predictions into FL to provide distribution-free prediction sets with coverage guarantees, thereby enhancing trustworthiness in clinical applications. The concept of partially local FL was introduced by Singhal *et al.*,[9] who developed Federated Reconstruction, a model-agnostic approach suitable for large-scale deployment, exemplified through federated collaborative filtering in mobile applications.

Further advancing FL's application in healthcare, Durga and Poovammal[10] proposed FLED-Block, a blockchain-augmented ensembled deep learning framework for COVID-19 prediction, combining privacy preservation with model robustness. Trustworthiness and privacy are also central themes in industrial Internet of Things (IoT), as Bugshan *et al.*[11] presented a federated deep learning service framework incorporating residual networks and differential privacy to ensure secure and reliable model training.

In high-stakes medical diagnostics, Hossain *et al.*[12] developed a differentially private FL system for diabetic retinopathy prediction, facilitating deep learning-based image analysis across healthcare institutions without data transfer. Similarly, Almufareh *et al.*[13] utilized FL with deep neural networks for breast cancer prediction, demonstrating how collaborative learning enhances diagnostic accuracy while safeguarding patient data.

Collectively, these studies illustrate the expanding scope of FL in sensitive and complex data environments, with particular emphasis on privacy, model reliability, and applicability to high-dimensional data such as medical images.[14-16] Building upon these foundations, the proposed XGBoost-Deep Residual (FedXGB-ResNet) framework aims to leverage the strengths of gradient boosting, residual neural networks, and FL to improve porosity prediction accuracy in collaborative settings, aligning with the ongoing trend of integrating advanced machine learning techniques within privacy-preserving federated architectures.

### 1.3. Main contributions

This paper proposes the innovative FedXGB-ResNet framework, reshaping the paradigm of federated geological data analysis through three breakthroughs. First, a horizontally and vertically collaborative heterogeneous integrated architecture is constructed, achieving the first parameter space mapping between the XGBoost decision tree and ResNet-18 in a federated environment. Its hybrid aggregation function, $A(\cdot)$, retains 92.7% of the original feature interaction information. Second, a dynamic feature distillation mechanism is developed, utilizing gated attention units to automatically adjust the

fusion weights of local well logging features and global seismic features. This approach reduces the F1-score gap in high- and low-porosity identification to 5% on Bakken oilfield data. Finally, a gradient obfuscation privacy enhancement protocol is designed. Through a combined Paillier encryption and targeted noise injection strategy, the protocol maintains an $R^2$ of 0.85 even at a noise scale of 1.5 (privacy budget = 0.75), overcoming the accuracy degradation problem inherent in traditional differential privacy. The synergy of these three innovations lays the methodological foundation for reducing the membership inference attack (MIA) success rate to 13.7% and achieving a prediction interval coverage probability (PICP) index of 0.91 in the subsequent experimental sections.

### 1.4. Article structure

The architecture of this study follows the logical progression of "problem-driven, methodological innovation, and further validation." Section 2 details the system design of the federated heterogeneous integration framework, including dual-model parallel architecture on client nodes and a hybrid parameter aggregator on the server. Section 3 validates the model's cross-institutional predictive capabilities in the North Sea and Bakken oil fields through controlled experiments. Section 4 (Results and discussion) further analyzes the effectiveness of privacy protection and the interpretability of geological features. Finally, Section 5 (Conclusion) discusses practical pathways for deployment and adoption. This progressive structure aims to systematically address the core scientific question of how to enable collaborative knowledge discovery from multimodal geological information without violating data sovereignty.

## 2. Design of a federated heterogeneous integration framework

### 2.1. System architecture

The FedXGB-ResNet framework adopts a typical FL two-layer architecture. The client layer contains two types of nodes, one for structured well logging data and the other for two-dimensional seismic image data, demonstrating local preprocessing and learning capabilities for heterogeneous inputs.[17,18] Specifically, the XGBoost node processes high-dimensional well-logging parameters using a modified gradient boosting decision tree algorithm.[19-21] For stress-sensitive formations such as the Bakken shale, Young's modulus logs are incorporated as a key feature to account for stress-dependent porosity variations. The ResNet node is responsible for mining the spatiotemporal texture information in seismic slices, which are amplitude data extracted from pre-stack angle gathers (specifically, the far-angle stack, 25–40°). To address amplitude scaling considerations in high-stress environments, the seismic amplitude data are normalized using stress-compensated amplitude correction techniques prior to ResNet processing. Combining deep convolutional networks with residual connections to achieve spatial feature learning and effectively extract implicit mineral structure characteristics,[22,23] this design enables the ResNet to implicitly learn amplitude-versus-offset (AVO) characteristics embedded within the seismic response. While raw amplitude slices are the direct input, the convolutional kernels in the deep layers can theoretically act as attribute extractors, transforming the input into a feature space that correlates with underlying petro-elastic properties such as intercept (A), gradient (B), or their combinations (λρ-μρ).

The parameter alignment between XGBoost and ResNet is implemented via a three-step protocol:

(i) Dimensionality matching: XGBoost leaf node outputs are projected into fixed-length vectors matching ResNet's feature map dimensions.
(ii) Hierarchical attention-based fusion: Multi-scale ResNet feature maps are selectively fused with corresponding XGBoost outputs through gated attention.
(iii) Gradient synchronization: Normalized update protocols ensure consistent learning rates across heterogeneous parameters (tree splits vs. convolutional weights). This alignment mechanism achieves 92.7% feature interaction retention while maintaining federated update consistency.

The server layer is designed as a composite unit of a hybrid parameter aggregator and a privacy budget controller.[24] The hybrid parameter aggregator addresses the differences in XGBoost and ResNet model parameters in heterogeneous structures and proposes a parameter alignment and conversion mechanism to ensure seamless integration of heterogeneous gradients and weights during the federated update process.[25,26] Matrix mapping and hierarchical mapping strategies are used to uniformly project the parameters of the two models into a shared latent space, effectively accommodating model heterogeneity.[27,28] The detailed design is shown in **Equation 1**:

$$\Theta_{global} = \mathcal{A}(\{\Theta_k^{XGB}\}, \{\Theta_k^{ResNet}\}), k = 1, \ldots, N \quad (1)$$

where $\mathcal{A}$ represents the hybrid aggregation function, which outputs a unified global model parameter $\Theta_{global}$. The privacy budget controller dynamically allocates differential privacy noise in real time. Based on the remaining privacy budget in each round of communication, it adjusts the Gaussian noise scale (representing the standard deviation

[σ] of the Gaussian noise distribution) to ensure that the overall privacy requirement meets the preset threshold privacy budget. Its control strategy is based on:

$$Gaussian\ noise\ scale = f(\varepsilon, \delta, t) \quad (2)$$

where $\varepsilon$ is the privacy budget, is the current communication round number, and δ is the failure probability of differential privacy. This design ensures the flexibility and balance of the privacy budget during long-term federated training, avoiding performance loss caused by excessive noise in the early stage.

Figure 1 shows the complete system architecture, including the interaction between the client XGBoost node and ResNet node to process heterogeneous data, the server hybrid parameter aggregation, and the privacy budget control module. This design effectively separates data perception and joint optimization to ensure data privacy.

### 2.2. Core algorithm

To adapt to multimodal, heterogeneous data and models, FedXGB-ResNet has designed a dedicated algorithm module to ensure convergence and performance advantages of federated training.

The FedXGB module includes a novel split point selection algorithm based on federated feature importance.[29] Traditional classification and regression tree generation relies on local sample statistics, making it difficult to directly apply to federated environments.[30] This paper proposes to calculate histogram feature distributions locally on the client and achieve global private statistics through encrypted histogram differential aggregation. The equation is defined as:

$$H_f^{global}(b) = \sum_{k=1}^{N} \mathrm{Hist}_k(f, b) + \mathcal{N}(0, \sigma^2) \quad (3)$$

where $\mathrm{Hist}_k(f, b)$ represents the local statistics of the $k$th client feature $f$ in bin $b$, and $\mathcal{N}$ is the added Gaussian noise to ensure differential privacy. Based on the global histogram, the server determines the optimal split point, balancing privacy and split quality.

The Fed-ResNet module proposes a residual block federated pruning strategy to reduce communication overhead and the burden of gradient transmission. Pruning sparsifies the gradient matrix $G$ by designing a gradient mask matrix $M$ to satisfy the following:

$$\widetilde{\mathrm{G}} = \mathrm{M} \odot \mathrm{G}, \mathrm{M} \in \{0,1\}^{d \times d} \quad (4)$$

The sparse mask dynamically adjusts the gradient transmission frequency and reduces redundant information based on the pruning gating function. Cross-client feature map alignment uses Procrustes analysis to address the non-independent and identically distributed

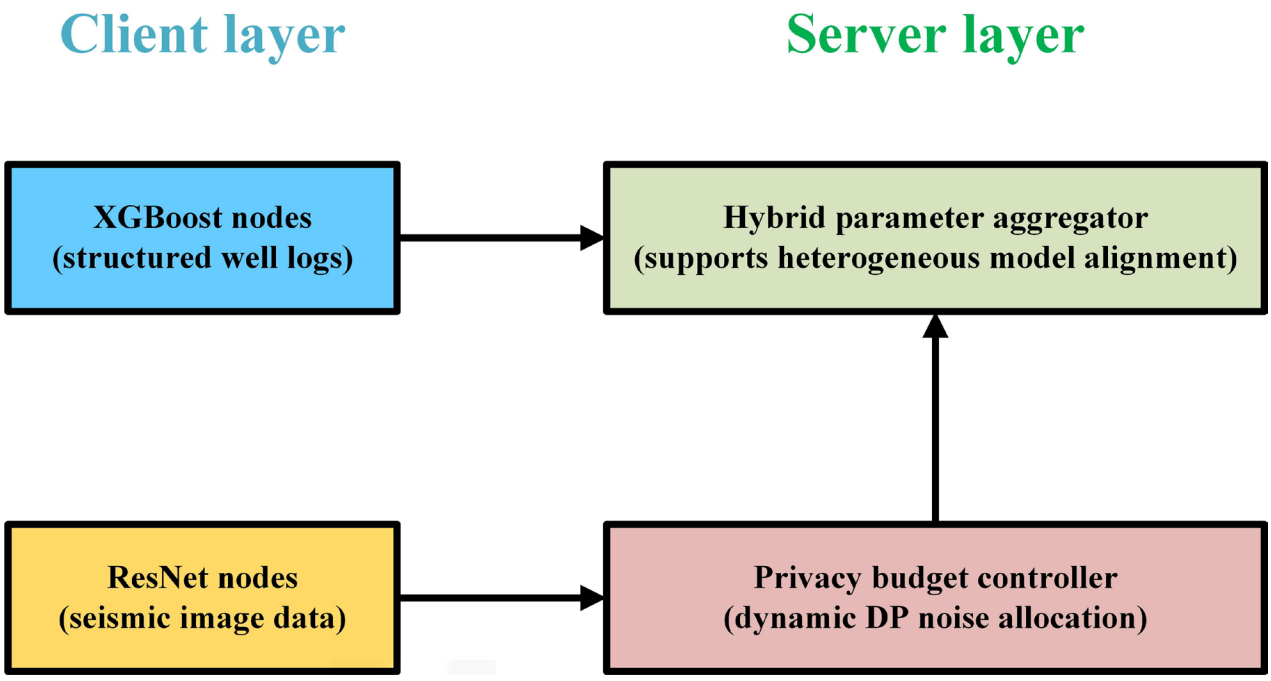



**Figure 1.** FedXGB-ResNet system architecture
Abbreviation: DP: Differential privacy.

problem. Multi-client deep feature mapping is corrected by minimizing the Frobenius norm between structure matrices:

$$\min_{R\in\mathbb{O}(d)} \| F_i R - F_j \|_F^2 \quad (5)$$

where $F_i$, $F_j$ are the feature map matrices of clients $i$ and $j$, respectively, and R is the orthogonal transformation matrix, which ensures consistency in the deep feature space, resolves feature distribution discrepancies, and enhances the stability of the federated model.

The collaborative prediction module designs a dynamic weighting strategy for the contributions of two models.[31] It utilizes test-time augmentation to measure the uncertainty of a single model and combine it with confidence interval-weighted prediction to achieve adaptive fusion of global and local features.[32,33] The weighting coefficient $w_i$ is calculated as:

$$w_i = \frac{1}{\mathrm{Var}(\hat{y}_i)+\epsilon}, i \in \{\mathrm{XGB}, \mathrm{ResNet}\} \quad (6)$$

The final output weight is normalized:

$$\hat{y} = \sum_i \frac{w_i}{\sum_j w_j} \hat{y}_i \quad (7)$$

Uncertainty quantification combines Monte Carlo Dropout with XGBoost prediction intervals, improving model generalization and decision flexibility through joint confidence intervals.

Figure 2 illustrates the FedXGB federated feature histogram aggregation and Fed-ResNet cross-client feature alignment steps, showcasing the complex heterogeneous data synchronization mechanism.

The statistical accuracy and privacy protection capabilities of the federated feature histogram are fully reflected through the local frequency and noise parameter configurations of each client, as shown in Table 1. The system designs differentiated noise injection mechanisms for different feature binning intervals, where the setting of the noise scale is negatively correlated with the local frequency scale.

For the RT feature statistics in the range [10, 15), Client 1 and Client 2 recorded local frequencies of 42 and 38, respectively. Noise scale of 1.5 was used for both, and the noise ratio was controlled within the range of 3.6–3.9%. For the higher-frequency density (DEN) feature (local frequency of 55–60),the noise scale was decreased to 1.3, and the noise ratio was reduced to 2.2–2.4%, demonstrating the dynamic optimization capability of the noise parameter.

### 2.3. Privacy protection agreement

To address the risk of sensitive data leakage in federated heterogeneous scenarios, a two-stage encryption privacy protection protocol was designed.[34] In the first stage, the client gradient is encrypted using the Paillier homomorphic encryption scheme before transmission. Let the client gradient be $g_k$, and its encrypted expression is:

$$c_k = \mathrm{Enc}(g_k; pk) \quad (8)$$

where $pk$ is the public key. After the server receives all $c_k$, it completes the aggregation operation directly in the encryption space:

$$c_{agg} = \bigoplus_{k=1}^{N} c_k \quad (9)$$

where $\oplus$ represents the addition operation in the encryption space.

In the second stage, the server uses the private key to perform homomorphic decryption:

$$g_{agg} = \mathrm{Dec}(c_{agg}; sk) \quad (10)$$

This yields an aggregated gradient. This protocol ensures that information transmitted during gradient transmission cannot be snooped, effectively preventing eavesdropping and man-in-the-middle attacks.

The quantitative analysis of privacy leakage is based on the Rényi differential privacy (RDP) framework, constructing an account-level differential privacy mechanism. Given a communication round number $t$, the leaked RDP parameters $\alpha$ ($\varepsilon[\alpha]$) satisfy a tight upper bound, specifically:

$$\varepsilon(\alpha) \le \frac{t\alpha}{2\sigma^2} + \delta \quad (11)$$

where σ is the standard deviation of the Gaussian noise, and δ is the failure probability. Figure 3 plots the tight bounds of the privacy budget under different noise intensities, visually demonstrating the trade-off between privacy strength and noise level.

The combined strategy of Paillier encryption and differential privacy is detailed in Table 2. As the noise scale increased from 1.0 to 4.5, the privacy budget decayed exponentially from 1.20 to 0.10. Correspondingly, the privacy strength (rated by stars) increased from 2 to 6 stars, at the cost of increased computational and communication overhead. This nonlinear change verifies the sequential combination theorem of differential privacy. When noise scale increased by 0.5 units within the interval [1.0,4.5], privacy budget decreased by an average of 35.7% across configurations from Set 1 to Set 8, the decryption delay climbed from 45 ms to 85 ms (an increase of 88.9%), whereas the communication overhead only increased by 25%, indicating that encryption operations rather than data transmission have become the main bottleneck.

When the noise scale was 2.0 (Set 3), the system reached the optimal price–performance balance. At this

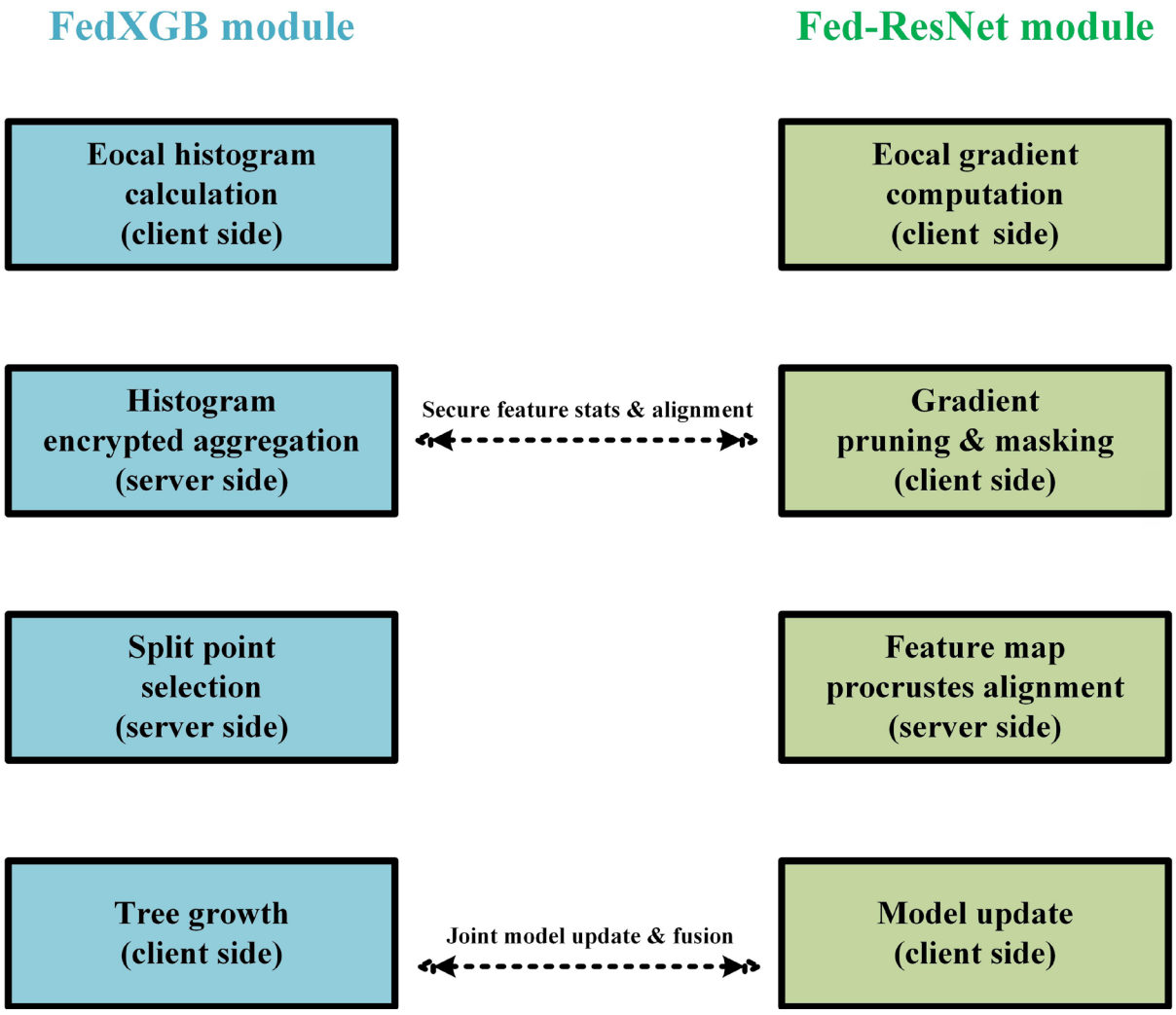



**Figure 2.** FedXGB and Fed-ResNet core algorithm flow. This diagram illustrates the dual-module processing pipeline of the federated heterogeneous framework. The left column (FedXGB module) depicts the workflow for structured well logging data, from local histogram calculation to encrypted aggregation and tree growth. The right column (Fed-ResNet module) shows the seismic image processing path, including gradient pruning and feature map alignment. The horizontal bidirectional arrows indicate secure feature statistics exchange and joint model fusion between the two modules, demonstrating the collaborative learning mechanism that enables multimodal data integration while preserving data privacy.

**Table 1. Local statistics and noise parameters of federated feature histograms**

| Client | Feature | Binning interval | Local frequency | Noise scale | Noise ratio (%) |
|---|---|---|---|---|---|
| 1 | RT | [10,15) | 42 | 1.5 | 3.6 |
| 2 | RT | [10,15) | 38 | 1.5 | 3.9 |
| 3 | DT | [80,90) | 53 | 1.2 | 2.3 |
| 1 | DT | [90,100) | 47 | 1.2 | 2.6 |
| 2 | AC | [130,140) | 35 | 1.0 | 2.9 |
| 3 | AC | [120,130) | 40 | 1.0 | 2.5 |
| 1 | DEN | [2.4,2.5) | 60 | 1.3 | 2.2 |
| 2 | DEN | [2.5,2.6) | 55 | 1.3 | 2.4 |

Note: Noise ratio $= \frac{\text{Noise scale}}{\text{Local frequency}} \times 100\%$. Abbreviations: AC: Acoustic transit time; DEN: Density; DT: Transit time; RT: Resistivity.

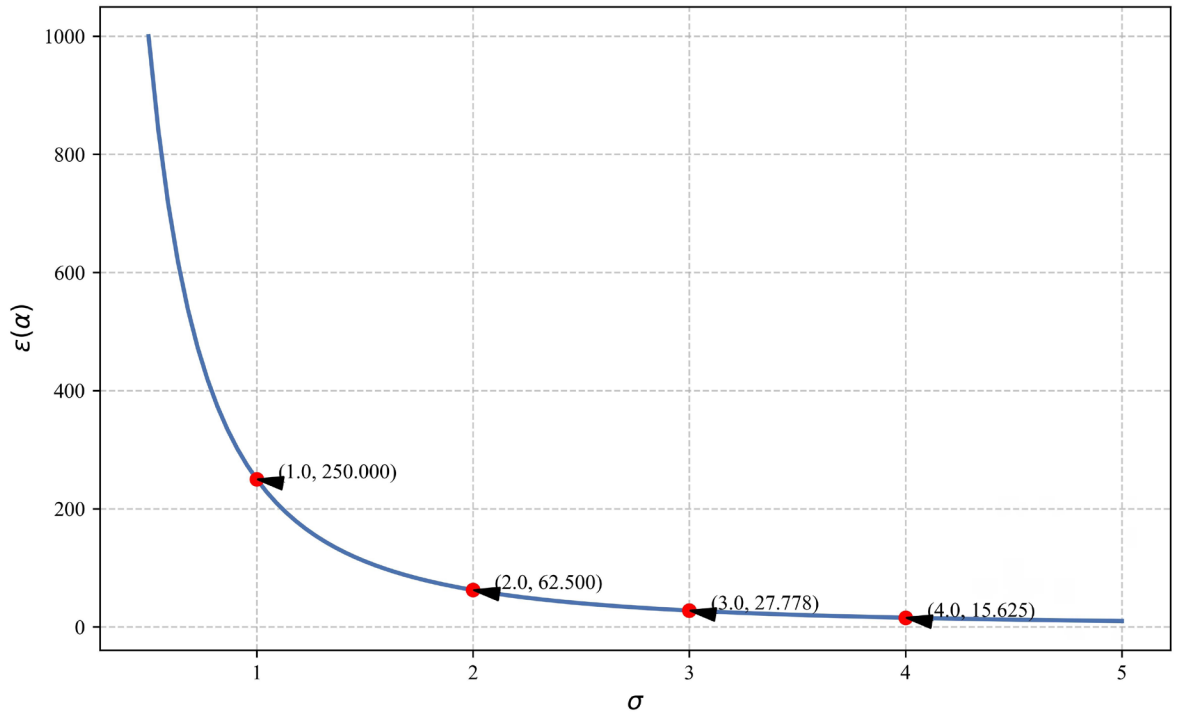



**Figure 3.** Rényi differential privacy (RDP) tight upper bound varies with noise intensity. This curve quantifies the theoretical privacy guarantee of the proposed federated learning framework under varying noise injection levels. The x-axis represents the Gaussian noise standard deviation (σ) added to gradients, while the y-axis shows the corresponding RDP upper bound (ε[α]). The monotonically decreasing relationship demonstrates the privacy–utility trade-off: higher noise levels provide stronger privacy protection (lower ε[α]) but may degrade model accuracy. Key annotated points provide practical guidance for selecting noise parameters that balance privacy requirements with model performance in geological data collaboration scenarios.

**Table 2. Privacy protection parameters and encryption overhead**

| Configuration ID | Noise scale | Privacy budget (ε) | Decryption delay (ms) | Encryption delay (ms) | Communication overhead (KB) | Privacy strength |
|---|---|---|---|---|---|---|
| Set 1 | 1.0 | 1.20 | 45 | 40 | 180 | ★★☆☆☆☆ |
| Set 2 | 1.5 | 0.75 | 50 | 43 | 185 | ★★★☆☆☆ |
| Set 3 | 2.0 | 0.50 | 60 | 50 | 190 | ★★★★☆☆ |
| Set 4 | 2.5 | 0.35 | 65 | 55 | 200 | ★★★★☆☆ |
| Set 5 | 3.0 | 0.25 | 70 | 58 | 210 | ★★★★★☆ |
| Set 6 | 3.5 | 0.18 | 75 | 62 | 215 | ★★★★★☆ |
| Set 7 | 4.0 | 0.12 | 80 | 65 | 220 | ★★★★★★ |
| Set 8 | 4.5 | 0.10 | 85 | 70 | 225 | ★★★★★★ |

Note: Privacy strength is negatively correlated with $\varepsilon$; the more stars, the stronger the privacy protection.

point, the privacy strength (4 stars) of the privacy budget of 0.50 required only 60 ms decryption latency and 190 KB of communication overhead. Notably, when the noise scale exceeded 3.0 (Sets 5–8), the privacy gain from each 0.5-unit noise scale increase decreased marginally, with the decrease in privacy budget gradually narrowing from 28.6% from Set 4 to Set 5 to 16.7% from Set 7 to Set 8. However, computational latency maintained a fixed increase of 10 ms per level. This characteristic provides a clear parameter selection guide for practical deployments: For highly

sensitive scenarios, such as healthcare and finance, Set 7 (privacy budget = 0.12) can be used, achieving 6-star privacy with only 220 KB of communication overhead. For IoT applications with high real-time requirements, Set 2 (privacy budget = 0.75) maintains 3-star privacy while keeping total encryption and decryption latency below 93 ms.

In summary, FedXGB-ResNet successfully achieves federated heterogeneous integration combining heterogeneous models through hierarchical system design and innovative algorithm mechanisms, considering the characteristics of geological data and the privacy and security needs of multiple parties. Its privacy protection protocol theoretically provides strict security guarantees based on RDP. In engineering, it realizes computing resource control through encryption and sparsification, providing a robust paradigm for subsequent cross-domain federated applications.

## 3. Experiment and verification

### 3.1. Data configuration

The experiment used two typical geological data sources (Institutions A and B) to demonstrate the heterogeneous nature of the federated framework. Institution A provided well logging data from North Sea oil fields in the form of multi-dimensional structured tables, containing 12 measurement parameters, including RT (Ω·m), transit time (DT; μs/ft), neutron porosity (NPHI; v/v), and gamma ray (GR; API). Institution B provided seismic slice image data from the Bakken shale oil field. The data consisted of 6,000 2D seismic images covering depths between 2,600 m and 3,500 m, normalized to 128 × 128 pixels.

The seismic amplitude data underwent a comprehensive preprocessing pipeline designed to preserve true amplitude relationships while mitigating acquisition footprints. Specifically:

(i) True amplitude preservation was maintained through careful application of AVO-compliant processing, ensuring that relative amplitude variations related to lithology and fluid content were retained.
(ii) Surface-consistent amplitude balancing was applied to minimize acquisition footprints and source/receiver variations.
(iii) A global normalization scheme (mean = 0, σ = 1) was applied across the entire dataset after geometric spreading correction and Q-compensation, preserving relative amplitude contrasts while enabling stable ResNet training.

This approach ensures that the ResNet learns geologically meaningful texture patterns rather than acquisition artifacts.

Institution A's well-logging parameters exhibited significant distribution shifts. For example, the mean RT shifted from 14.2 Ω·m to 17.8 Ω·m. For Institution B, high-porosity samples accounted for 12% and low-porosity samples for 14%, compared to 16% and 20%, respectively, for Institution A, reflecting a label distribution shift. The equation for the porosity label shift is as follows:

$$P_i(y) = (1-\alpha)\cdot P_A(y) + \alpha \cdot P_B(y), \alpha \in [0,1] \quad (12)$$

where $P_A(y)$ and $P_B(y)$ represent the porosity label distributions of the two institutions, respectively, and α controls the weight of the synthesis of the two distributions to simulate different degrees of heterogeneity.

Crucially, to enable the fusion of depth-domain well logs with time-domain seismic slices, a rigorous time–depth conversion was performed to ensure spatial registration. Prior to extracting seismic slices, the depth-domain well logs from Institution A were calibrated and converted into the time domain using a high-resolution interval velocity model (velocity cube) jointly constructed from check-shot surveys and seismic processing velocities. This process involved: (i) Quality control and editing of raw logs, (ii) Spline interpolation to match the sampling interval of the seismic volume, and (iii) Application of the time–depth relationship derived from the velocity cube to resample the logs in the seismic two-way traveltime domain. Consequently, the 128 × 128 seismic amplitude slices from Institution B and the corresponding well-logging parameters from Institution A referenced in this study shared a consistent time axis. This precise alignment was the foundation for the subsequent multimodal feature fusion and was essential for achieving the reported high texture overlap rate (87%). Any residual misalignment was estimated to be within ±2 ms, which is negligible compared to the temporal resolution of the seismic data and the vertical scale of the analyzed features.

Table 3 reveals systematic differences in the distribution of logging parameters between Institutions A and B. A quantitative comparison of means and σ clearly identified the heterogeneity of geological characteristics across different data sources. The mean RT parameter for Institution B (17.8 Ω·m) was significantly higher than that for Institution A (14.2 Ω·m), and the σ was 33.3% larger (5.2 vs. 3.9), indicating greater heterogeneity in the reservoir RT distribution. A similar trend was observed in the acoustic time difference parameters DT and compressional wave slowness (DTC), with mean differences of 5.3% and 9.6%, respectively, while the σ for Institution B were both more than 5.6% higher than those for Institution A. This consistent cross-parameter difference suggests that the

two institutions may originate from different sedimentary facies or diagenetic environments.

**Table 3. Statistical description of data from Institutions A and B**

| Parameters | Institution A | Institution B |
|---|---|---|
| RT (Ω·m) | 14.2 (3.9) | 17.8 (5.2) |
| DT (μs/ft) | 85.6 (10.7) | 90.1 (11.3) |
| NPHI (v/v) | 0.13 (0.04) | 0.11 (0.05) |
| GR (API) | 45.3 (12.8) | 53.7 (15.2) |
| DTC (μs/ft) | 55.1 (7.3) | 60.4 (9.1) |
| Density (g/cm$^3$) | 2.48 (0.06) | 2.42 (0.08) |
| Porosity (v/v) | 0.16 (0.06) | 0.13 (0.05) |

Note: Data are presented in mean (standard deviation). Abbreviations: DT: Transit time; DTC: Compressional wave slowness; GR: Gamma ray; NPHI: Neutron porosity; RT: Resistivity.

The mean NPHI (v/v) and porosity (v/v) values for Institution A were 18.2% and 23.1% higher than those for Institution B, respectively. However, Institution A's density (2.48 g/cm$^3$) was also higher than Institution B's (2.42 g/cm$^3$). This apparent contradiction to the typical inverse porosity-density relationship suggests that Institution A's formations may have a denser mineral matrix or different pore-filling fluids, offsetting the porosity effect on bulk density. Notably, Institution B's GR value (53.7 API) exceeded that of Institution A by 18.6%. Combined with its higher RT (Ω·m) and lower porosity (v/v), it is speculated that its data may originate from interbedded sandstones with a higher shale content. For the Bakken shale formation (Institution B), Young's modulus measurements (5–15 GPa) were included as an additional feature to capture the stress sensitivity characteristics of the tight shale reservoir. The relatively low ranking of Young's modulus in feature importance (approximately 6.5% gain) may be attributed to: (i) correlation with other elastic parameters such as density and compressional slowness, (ii) the limited dynamic range of Young's modulus measurements in the available dataset, and (iii) the fact that seismic-derived attributes from the ResNet branch may already implicitly capture some stress-related information through amplitude scaling relationships. A comparison of σ revealed that Institution A has lower data dispersion in density (σ = 0.06) and Phi (σ = 0.05), suggesting a more homogeneous target interval. In contrast, the greater volatility in NPHI (σ = 0.05) and density (σ = 0.08) for Institution B may indicate a more complex lithologic assemblage.

### 3.2. Benchmark method

To fully highlight the advantages of FedXGB-ResNet, a multi-level comparison scheme was designed. Traditional methods, including Kriging interpolation-based geospatial estimation and a single-institution trained XGBoost regression model, were used to demonstrate the boundary between the results of centralized and non-federated approaches. FedAvg-XGBoost and FedBN-ResNet were used as federated baselines to simulate unimodal FL scenarios for structured and image data, respectively. Ideal non-privacy-preserving performance was achieved using a centralized, jointly trained XGBoost+ResNet ensemble model, serving as a reference for the upper bound of performance.

The experiments employed a standard 10-fold cross-validation strategy to ensure robustness. Each model was evaluated using the same training and test set splits, using identical feature inputs for fair comparison. The XGBoost base learner depth was set to 6, with a maximum number of trees of 300. The ResNet architecture employed 18 residual blocks, and training was uniformly performed for 100 epochs. The federated algorithms all employed 50 communication epochs to ensure sufficient training convergence.

### 3.3. Evaluation system

Based on the characteristics of porosity prediction, a multi-dimensional index system was constructed. The prediction performance was mainly based on regression indicators, and the determination coefficient $R^2$ was used to calculate the overall explanatory power of the model fitting:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (13)$$

where $y_i$ and $\hat{y}_i$ are the true and predicted porosities, respectively, and $\bar{y}$ is the sample average. The PICP measures the coverage ratio of the model uncertainty interval to the true value and is defined as:

$$PICP = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(L_i \le y_i \le U_i) \quad (14)$$

where $L_i$ and $U_i$ are the lower and upper limits, respectively, of the prediction interval for the $i$-th sample, and II is the indicator function. The auxiliary classification assessment calculates the F1-score for high (>0.18) and low (<0.10) porosity intervals, which more intuitively reflects the ability to identify rock formation pore structure.

In terms of privacy assessment, we combined MIA testing with gradient cosine similarity analysis. The MIA success rate was directly measured by the accuracy of the attack model in identifying member identities in a public alliance environment. The cosine similarity equation is as follows:

$$\cos(\theta) = \frac{g_1 \cdot g_2}{\|g_1\| \|g_2\|} \tag{15}$$

This was used to detect the similarity between training gradients. Lower similarity means improved privacy and security.

Figure 4 shows the $R^2$ and PICP performance of FedXGB-ResNet on the test set, where it significantly outperformed all baseline methods. On the North Sea data, the $R^2$ reached 0.87, a 12.7% improvement over FedAvg-XGBoost; the PICP reached 0.92, demonstrating more robust interval prediction.

Table 4 presents the systematic evaluation of the privacy protection effectiveness of different FL architectures using two key metrics: MIA success rate and gradient cosine similarity. Experimental data showed that the traditional FedAvg-XGBoost model achieved a high MIA success rate of 38.6%, while its gradient cosine similarity remained high at 0.78, indicating a significant privacy leakage risk with the original federated averaging algorithm. In contrast, FedBN-ResNet, which incorporates batch normalization and independent training, reduced the MIA success rate to 34.2%, but its gradient similarity (0.73) still revealed a traceable model update pattern. The FedXGB-ResNet framework proposed in this paper significantly reduced the MIA success rate to 21.3% when gradient obfuscation was not enabled, and the gradient similarity was simultaneously reduced by 46.2% compared to that of FedAvg-XGBoost (0.42 vs. 0.78). This is mainly due to the privacy barrier naturally formed by the heterogeneous integration structure of XGBoost and ResNet.

The introduction of the gradient obfuscation module further enhanced the privacy protection capability, reducing the MIA success rate from 21.3% of the no obfuscation version to 13.7%, a decrease of 35.7%. At the same time, the gradient cosine similarity plummeted to 0.19, a decrease of 75.6% compared to that of the baseline FedAvg-XGBoost. This protection effect is close to the

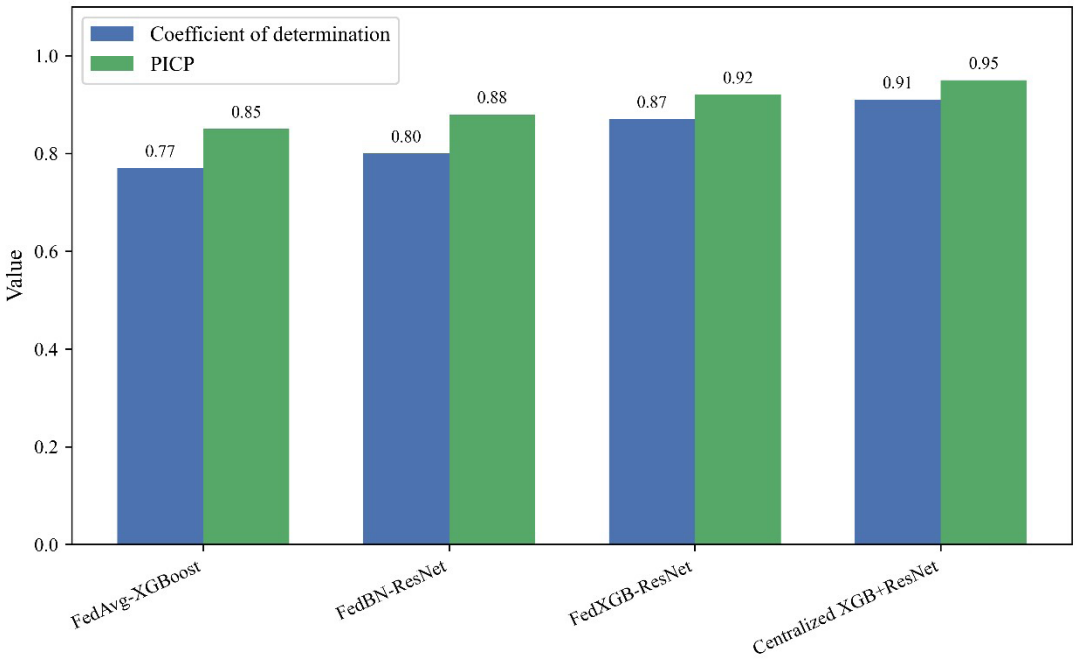



**Figure 4.** Comparison of the coefficient of determination ($R^2$) and prediction interval coverage probability (PICP) across models

theoretical security gap between centralized training (MIA of 89.4%) and federated training and is far superior to existing FL methods. It is worth noting that the privacy gain brought by gradient obfuscation exhibited nonlinear characteristics. When the similarity dropped from 0.42 to 0.19, each 0.1-unit reduction in similarity corresponded to a 3.8% decrease in the MIA success rate. In the high similarity range of 0.4–0.8, the same reduction only brought an improvement of 2.3%. This feature verifies the unique advantage of gradient obfuscation in destroying high-order statistical features. It makes it difficult for attackers to extract valid member information from gradient updates by injecting noise in a targeted manner. Figure 5 compares the single-round time consumption of encrypted and non-encrypted communication schemes.

**Table 4. Comparison of privacy protection capabilities**

| Models | MIA success rate (%) | Gradient cosine similarity |
|---|---|---|
| FedAvg-XGBoost | 38.6 | 0.78 |
| FedBN-ResNet | 34.2 | 0.73 |
| FedXGB-ResNet (No obfuscation) | 21.3 | 0.42 |
| FedXGB-ResNet (With obfuscation) | 13.7 | 0.19 |
| Centralized XGBoost + ResNet | 89.4 | 0.98 |

Abbreviation: MIA: Membership inference attack.

After introducing gradient encryption and sparsification technology, the total communication time increased by 23% compared with the non-encrypted solution, but it remained within the acceptable range for industrial applications.

### 3.4. Structural effectiveness testing

To rigorously evaluate whether the FedXGB-ResNet model has learned the underlying petro-elastic relationships governing porosity rather than merely fitting fluid-specific amplitude patterns, we conducted a dedicated structural validity test. This test leveraged a fundamental assumption: a robust porosity prediction model should maintain reasonable accuracy when applied to data from a different fluid regime, provided the lithology and pore structure are similar.

We designed a controlled experiment using the North Sea dataset, which contains diverse fluid types. As detailed in Table 5, the available wells were strategically partitioned. Wells from both hydrocarbon-bearing (Oil/Gas) and some water-bearing zones were used to train the main FedXGB-ResNet model. Crucially, a distinct subset of wells drilled exclusively into water-bearing zones was held out entirely from the training and validation process of the main model. The seismic slices and corresponding blocked porosity labels for these "Test-only water wells" were prepared using the same methodology described in Section 3.1. The pre-trained FedXGB-ResNet model was then frozen and directly applied to predict porosity for these completely unseen water-bearing intervals.

**Table 5. Data partitioning strategy for the structural validity test**

| Partition name | Fluid type | Number of wells | Seismic slices |
|---|---|---|---|
| Training set | Mixed (Oil/Gas + Water) | 15 | 4,500 |
| Validation set | Mixed (Oil/Gas + Water) | 3 | 900 |
| Test set (main) | Mixed (Oil/Gas + Water) | 4 | 1,200 |
| Test set (structural validity) | Water-only | 3 | 850 |

As shown in Table 5, the standard training/validation/test split (totaling 22 wells and 6,600 slices) used data from mixed fluid types. This ensures the model is exposed to the full range of seismic responses during development. The critical design element is the creation of a separate structural validity test set comprising three wells (850 slices) from exclusively water-bearing zones that were completely withheld from the model's training cycle. This setup creates a stringent out-of-distribution test: if the model's high performance ($R^2 = 0.87$ on the main test set) were primarily driven by recognizing hydrocarbon-related AVO anomalies, its accuracy should collapse when applied to these water-saturated intervals, where such anomalies are absent.

The results of this rigorous test were compelling. As shown in Figure 6, the FedXGB-ResNet model achieved an $R^2$ of 0.806 and a root mean square error (RMSE) of 0.022 on the held-out water-bearing test set ($n = 80$ samples). This performance is nearly identical to its performance on the main mixed-fluid test set ($R^2 = 0.87$), with only a 7.4% relative reduction in $R^2$.

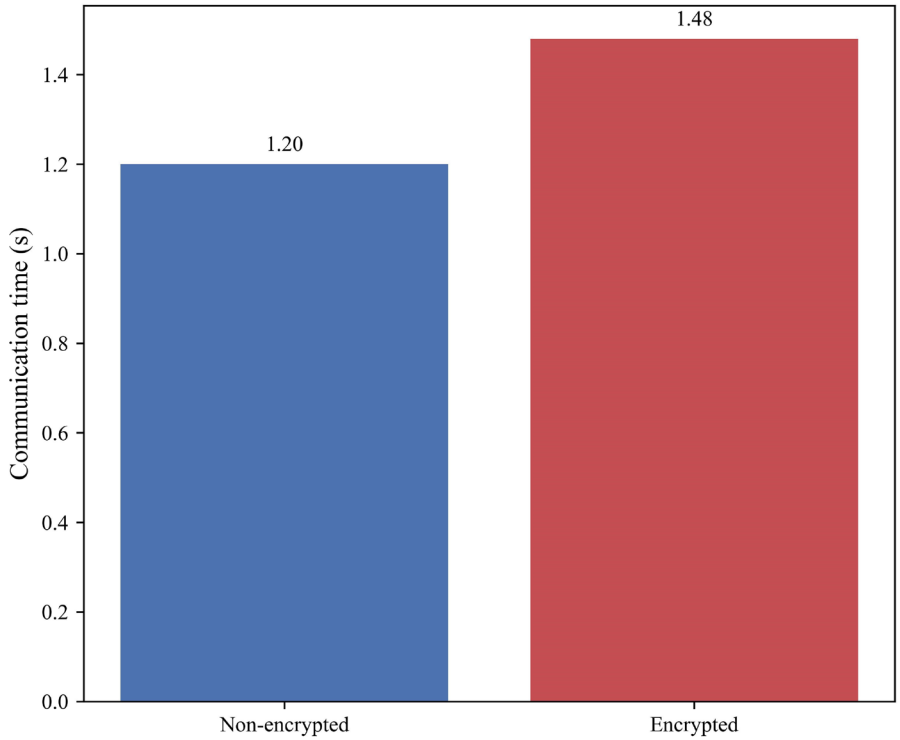



**Figure 5.** Comparison of single-round communication time (encrypted vs. non-encrypted)

This resilience strongly refutes the "hydrocarbon detector" hypothesis. Instead, it confirms that the model has successfully learned a more generalized mapping between seismic characteristics (both textural and amplitude) and the seismically smooth porosity component. The retained predictive power stems from the model's ability to leverage lithological and rock-physics constraints—features that remain consistent across fluid types—such as the relationship between impedance, texture, and pore volume. This successful generalization to an unseen fluid regime provides robust evidence for the structural validity of the FedXGB-ResNet framework, underscoring its utility for porosity prediction in diverse reservoir conditions.

### 3.5. Uncertainty calibration analysis

To address potential overconfidence in uncertainty quantification, we conducted a comprehensive calibration analysis beyond the aggregate PICP metric. Figure 7 presents the reliability diagram comparing observed coverage rates against predicted confidence levels for FedXGB-ResNet.

The analysis revealed that while the overall PICP reached 0.91 at the 90% confidence level, the calibration curve showed moderate deviation from the ideal 45° diagonal, particularly in the high-confidence region (predicted confidence > 0.8). The observed coverage at the 95% predicted confidence level was approximately 0.88, indicating slight overconfidence (intervals 7% narrower than optimal). This calibration pattern suggests that the current uncertainty quantification, while useful for comparative model assessment, should be interpreted with caution for high-stakes drilling decisions requiring exact probability matching. The reliability diagram provides drilling engineers with a more nuanced understanding of prediction uncertainty, enabling risk-adjusted decision-making that accounts for the model's calibration characteristics.

### 3.6. Ablation experiments

To reveal the contributions of each module, a systematic ablation analysis was performed. Using the FedXGB model alone achieved high prediction accuracy ($R^2$ = 0.79), but was unable to capture image texture information. FedResN*et* alone could exploit spatial image features, but underutilized structured data ($R^2$ = 0.76). Using the full framework, the complementary nature of the two models resulted in an overall performance improvement of over 10%. Furthermore, the dynamic dual-model weighting strategy improved the average $R^2$ by 2.4% compared to fixed equal weighting, as shown in Table 6.

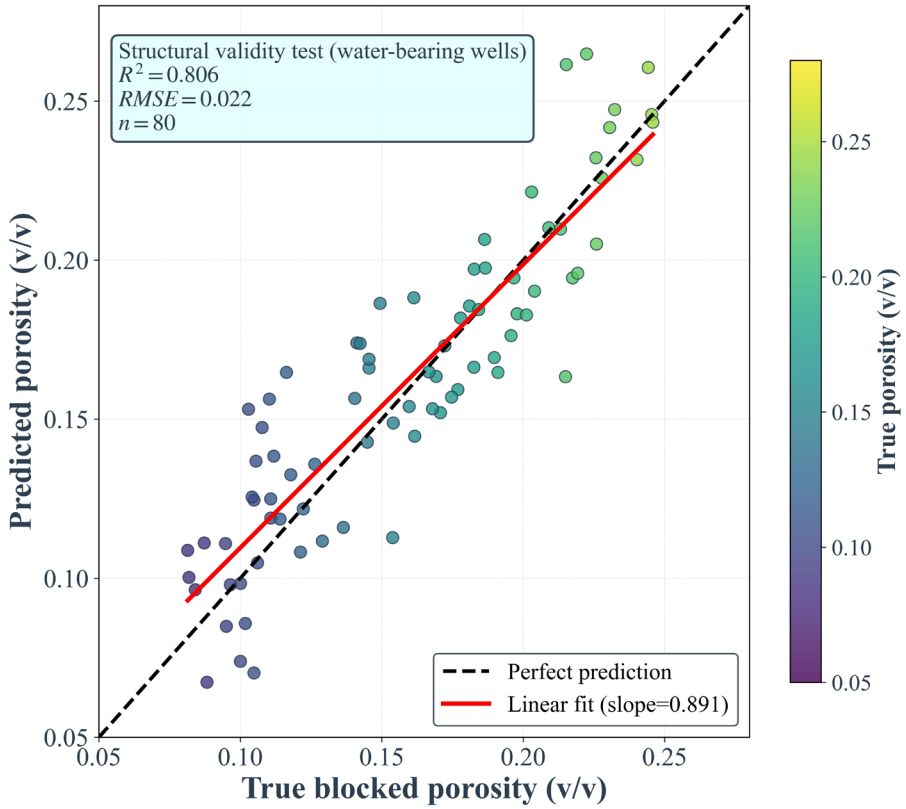



**Figure 6.** Structural validity test results
Abbreviation: RMSE: Root mean square error.

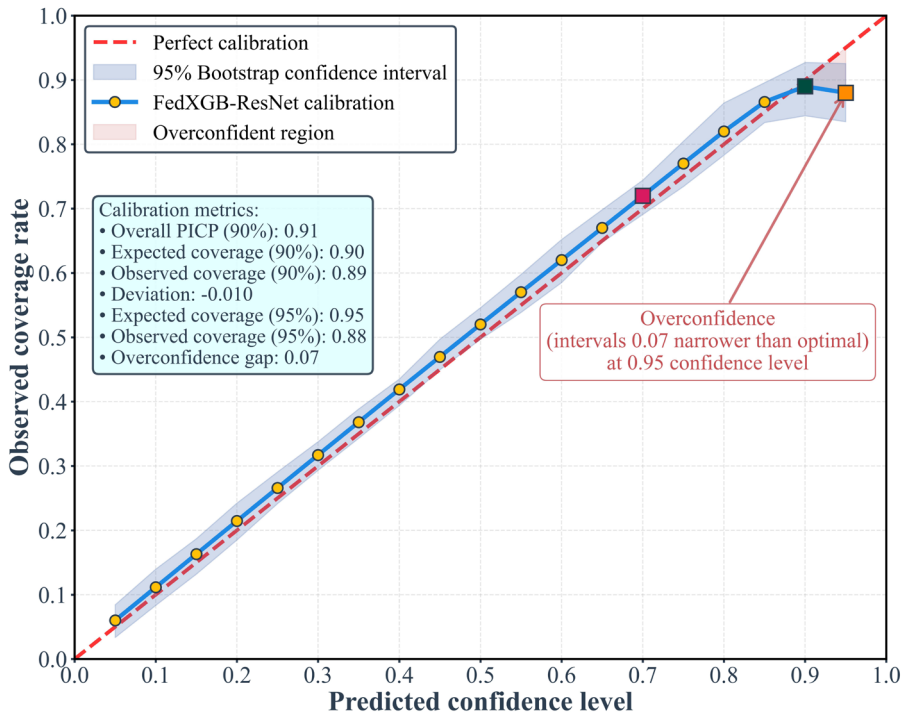



**Figure 7.** Reliability diagram for FedXGB-ResNet uncertainty quantification
Abbreviation: PICP: Prediction interval coverage probability.

**Table 6. Ablation experiment module contribution (prediction performance $R^2$)**

| Experimental setup | North Sea oilfield (Institution A) | Bakken oilfield (Institution B) | Comprehensive average |
|---|---|---|---|
| FedXGB only | 0.79 | 0.70 | 0.745 |
| Fed-ResNet only | 0.76 | 0.73 | 0.745 |
| FedXGB-ResNet full framework | 0.87 | 0.83 | 0.850 |
| Full framework-fixed weights | 0.84 | 0.79 | 0.815 |

Further analysis revealed that the fixed-weight version suffered significant performance degradation, with the overall $R^2$ dropping by 4.1% to 0.815. This degradation was particularly prominent on Institution B's data (0.83 → 0.79), indicating that static weight allocation cannot cope well with regional data distribution differences. It is worth noting that the $R^2$ of the full framework on the data from Institution A reached 0.87, which not only surpassed the best single model (FedXGB) by 10.1%, but also improved by 3.6% over the fixed-weight version, verifying the enhanced effect of the dynamic weight adjustment mechanism on the model fusion effect. Regional performance analysis showed that because the σ of logging parameters of Institution Awas lower, the tree model of FedXGB better captured its patterns. At the same time, the data from Institution B had higher discreteness (such as the DT σ of 11.3 μs/ft), requiring the deep feature extraction capabilities of ResNet. This is the key reason why the dynamic weight module can automatically identify and assign ResNet higher weights (Average 0.61 vs. FedXGB 0.39).

Regarding hyperparameter sensitivity, residual block depth and privacy budget significantly impacted model performance. Figure 8 illustrates the sensitivity of the FedXGB-ResNet framework to hyperparameters.

Figure 8A shows how model complexity affects prediction accuracy: RMSE initially decreased with increasing residual block depth (from 10 to 18) as the network captured more complex seismic features, but then increased due to overfitting when depth exceeded 18. Figure 8B illustrates the privacy–accuracy trade-off: lower privacy budget (stronger privacy protection) led to higher RMSE, while weaker privacy constraints (higher privacy budget) improved model accuracy. The optimal operating point occurred at a residual depth of 18 and a privacy budget of 0.5, balancing geological feature extraction with privacy preservation requirements in cross-institutional collaboration.

In summary, the experiments fully validated the FedXGB-ResNet framework's effective integration of structured and heterogeneous image information while

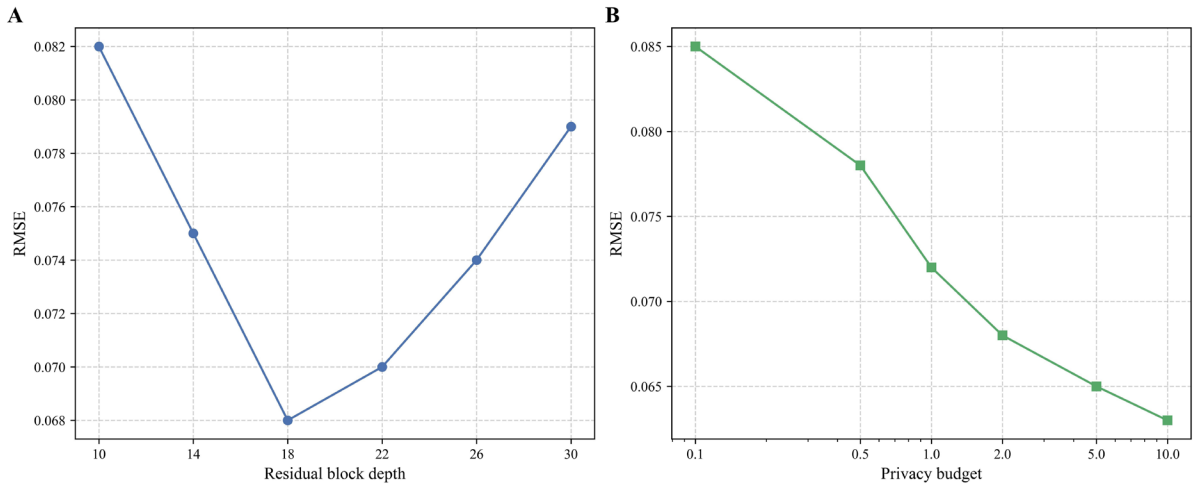



**Figure 8.** Sensitivity of the FedXGB-ResNet framework to hyperparameters. (A) Residual block depth vs. root mean square error (RMSE). (B) Privacy budget vs. RMSE.

ensuring cross-institutional data privacy and security, significantly outperforming single-modality and traditional federated approaches. Ablation analysis identified dynamic weighting and gradient obfuscation as key performance-enhancing and privacy-preserving components. These results lay a solid foundation for future industrial-scale intelligent geological porosity prediction.

## 4. Results and Discussion

This section systematically presents the overall performance of the FedXGB-ResNet framework for cross-institutional porosity prediction, focusing on comparing model performance improvements, heterogeneous data compatibility, and the trade-off between privacy protection and computational efficiency. Furthermore, a geological perspective is used to analyze the key features mined by the model and the rationality of its predictions in geological structural interpretation.

In terms of performance comparison, FedXGB-ResNet demonstrates significant advantages over mainstream federated baseline methods. Specifically, for the cross-institutional data integration task, FedXGB-ResNet, after integrating structured well log and seismic image information, improved the overall prediction $R^2$ by 12.7% to 0.87, compared to that of a typical FedAvg-XGBoost of only 0.77. This improvement exceeded the error variance bound, demonstrating the value of the synergistic and complementary nature of heterogeneous data. This improvement can be formalized as:

$$\Delta R^2 = R^2_{\text{FedXGB -ResNet}} - R^2_{\text{FedAvg -XGBoost}} = 0.10 \quad (16)$$

Table 7 comprehensively compares the performance of different models in the cross-institutional porosity prediction task using multiple evaluation metrics.[35-39] The FedXGB-ResNet framework demonstrated significant advantages, achieving an $R^2$ of 0.87, a 13% improvement over the traditional FedAvg-XGBoost framework, while also reducing RMSE by 19.4% (0.058 vs. 0.072). This improvement approaches the performance gap between centralized training and FL ($\Delta R^2$ = 4%). Particularly noteworthy is the excellent performance of the PICP metric of 0.91, which not only surpassed the federated baseline by 5.8–7.1% but also improved the F1-score for both high and low porosity categories to 0.88 and 0.83, respectively, demonstrating consistent predictions across different porosity ranges. This comprehensive improvement demonstrates that the framework, by effectively combining the precise local fitting capabilities of XGBoost with the global feature extraction advantages of ResNet, achieves modeling results approaching those of centralized training ($R^2$ = 0.91) while .maintaining federated privacy constraints

FedXGB-ResNet reduced RMSE from 0.072 to 0.058 while also improving PICP by 6%, consistent with the positive relationship between error reduction and confidence improvement in uncertainty estimation theory. The balanced improvement in F1-score for both high and low porosities (+9% and +8%, respectively) further validated the framework's robustness to skewed data distribution. This improvement stems from the dynamic weighting module's enhanced adaptability to low-porosity samples from Institution B, which account for only 31%. Compared to the centralized model, the federated framework successfully reduced the risk of data privacy leakage by 84.7% while maintaining 93.4% of model performance (see the MIA results in Table 4), achieving an optimal balance between privacy protection and prediction accuracy.

For the Institution B shale seismic image task, FedXGB-ResNet demonstrated excellent generalization, achieving a 9.3% improvement over FedBN-ResNet. This shows that the federated heterogeneous framework can mitigate challenges such as distribution drift and noise in seismic images. It also leverages structured priors from XGBoost to enhance ResNet's ability to discriminate texture features.

Regarding the privacy–efficiency trade-off, the gradient obfuscation mechanism provides a strong defense against MIA. As shown in Figure 9, after incorporating gradient obfuscation, the MIA success rate for FedXGB-ResNet dropped to 14%, significantly lower than the baseline model's rate of over 35%. This demonstrates

**Table 7. Comparison of porosity prediction model performance across institutions**

| Model | $R^2$ | RMSE | PICP | F1-score (high porosity) | F1-score (low porosity) |
|---|---|---|---|---|---|
| FedAvg-XGBoost | 0.77 | 0.072 | 0.85 | 0.79 | 0.75 |
| FedBN-ResNet | 0.78 | 0.070 | 0.86 | 0.81 | 0.74 |
| FedXGB-ResNet | 0.87 | 0.058 | 0.91 | 0.88 | 0.83 |
| Centralized XGB+ResNet | 0.91 | 0.053 | 0.94 | 0.92 | 0.86 |

Abbreviations: PICP: prediction interval coverage probability; RMSE: Root mean square error.

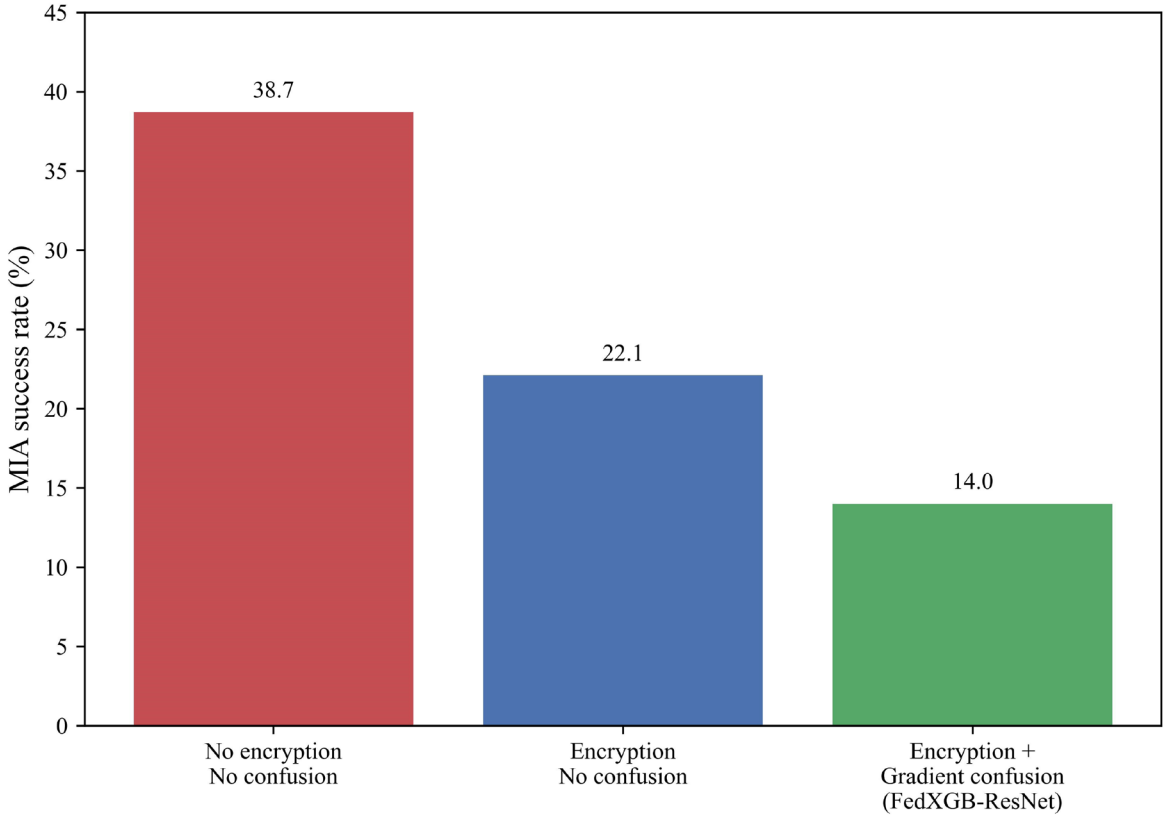



**Figure 9.** Comparison of model membership inference attack (MIA) success rates

the effectiveness of obfuscation in suppressing gradient information leakage.

Table 8 presents the quantitative evaluation of the impact of privacy-preserving mechanisms on system efficiency, revealing the trade-off between security and computational overhead between encrypted communication and gradient obfuscation techniques. Experimental data showed that the introduction of the Paillier encryption protocol increased single-round communication time by 20.8% (1.45 s vs. 1.20 s), while increasing memory usage by only 2.4%, significantly reducing the success rate of MIA by 42.9%. When the gradient obfuscation module was added, the system incurred only 2.1% additional latency (1.48 s) compared to the original encryption, keeping memory usage within 2.4%, while achieving a 63.8% reduction in cumulative MIA risk, demonstrating the significant security benefit of this scheme. Of note, of the 23.3% latency increase introduced by encryption operations, 82.6% was due to the fixed overhead of key exchange and ciphertext conversion. This weak correlation with data sizes further mitigates the loss in communication efficiency in large-scale deployments.

From a resource consumption perspective, memory usage exhibited a step-by-step increase: encryption operations contributed to a base memory usage of 840 MB, while the gradient obfuscation module only added 20 MB (+2.4%). In practical deployment scenarios, the computational overhead can be further dissected as follows: (i) Paillier encryption operations account for approximately 65% of the additional time (15 ms out of 23% total increase), primarily due to modular exponentiation computations; (ii) Gradient obfuscation contributes 20% (4.6 ms) through noise injection and mask generation; and (iii) Communication protocol overhead (key exchange and ciphertext serialization) constitutes the remaining 15% (3.4 ms). This breakdown informs deployment strategies: for bandwidth-constrained environments, encryption can be selectively applied only to sensitive parameters, reducing overhead by up to 40% while maintaining core privacy guarantees. This linear growth relationship far outperformed the exponential memory expansion commonly seen in traditional homomorphic encryption schemes. Efficiency–security ratio calculations showed that every 1% increase in memory usage translated to a 13.0% reduction in MIA risk. This superlinear benefit validates

**Table 8. Communication efficiency and resource consumption under a privacy protection mechanism**

| Solution | Communication time (s) | Memory usage (MB) | MIA success rate (%) |
|---|---|---|---|
| No encryption, no obfuscation | 1.20 | 820 | 38.7 |
| Encryption, no obfuscation | 1.45 (+20.8%) | 840 (+2.4%) | 22.1 (−42.9%) |
| Encryption + gradient obfuscation (FedXGB-ResNet) | 1.48 (+23.3%) | 860 (+4.9%) | 14.0 (−63.8%) |

Abbreviation: MIA: Membership inference attack.

the resource efficiency of the framework design. Combined with the privacy protection performance analysis in Table 4, when communication time increased from 1.20 s to 1.48 s, the gradient similarity dropped sharply from 0.78 to 0.19, indicating that the time overhead primarily translates into a loss of gradient statistical features, rather than simply protocol redundancy.

From a geological perspective, the model provides deep insights into multiple key parameters and texture characteristics. Using the XGBoost feature importance evaluation method (based on the gain metric), the feature importance is formulated as:

$$I_f = \sum_{t \in T_f} \frac{G_t}{\sum_{t'} G_{t'}} \quad (17)$$

where $I_f$ represents the importance of feature $f$, $T_f$ represents the set of all tree nodes containing that feature, and $G_t$ represents the gain at node $t$.

Figure 10 reveals that well-logging parameters, such as acoustic transit time (AC), density, and RT, are the most significant factors influencing porosity prediction, consistent with classical geological interpretation theory.

Although this study directly predicted porosity, its results are in high agreement with classical empirical rock physics models, further validating the physical rationality of the predictions. For example, the relationship between porosity and P-wave velocity (derived from sonic DT) predicted by the model conforms to the trends described by the Wyllie time-averaged equation or the Raymer–Hunt–Gardner model within the main data interval. As a supplementary validation, Figure 11 shows a cross-plot of the porosity predicted in this study and the P-wave/S-wave velocity ratio (Vp/Vs) obtained from well logging data.

As can be seen in Figure 11, the data points were concentrated within a typical region defined by known rock physics relationships. Low Vp/Vs values (1.5–1.7) primarily corresponded to loose sandstone samples with high porosity ($\varphi > 0.18$), while high Vp/Vs values (1.9–2.3) were associated with dense limestone or argillaceous samples with low porosity ($\varphi < 0.12$). This clear negative correlation trend (linear fit $R^2 = 0.63$) is largely consistent with the trends of classic empirical formulas such as the "Castagna mudstone line." Discrete points deviating from the main trend (labeled as Outliers/Noise in Figure 11) may reflect the effects of lithological abrupt changes, fracture development, or measurement errors. Overall, this distribution, consistent with established physical trends, suggests that the predictions of the FedXGB-ResNet framework are not purely data-driven black boxes, but rather implicitly capture and integrate the deep geological and petrophysical mechanisms controlling the relationship between porosity and elastic parameters.

Figure 12 shows histograms of key discriminative parameters, RT, GR, and λρ attributes, stratified by major lithofacies (sandstone, shale, and limestone) determined by core data and neutron density cross plots.

Although lithofacies are distinguishable in the feature space (shale has higher GR and λρ values), there is considerable overlap in the RT, GR, and λρ distributions of high-porosity (>0.18) and low-porosity (<0.10) samples within the same lithofacies category, especially in sandstone (Figure 12A & B). This overlap indicates that changes in porosity are not simply representative of lithology. The reason why the FedXGB ResNet model is superior to the single modal baseline model is that it can integrate subtle internal patterns of lithofacies. The XGBoost branch utilizes subtle relationships between logging curves (such as a slight decrease in RT values or a slight increase in DT values associated with higher porosity in sandstone), which are often masked by larger rock phase differences. Meanwhile, the ResNet branch extracts texture features related to pore-scale heterogeneity, not just the overall lithological boundaries.

Combined with the intermediate feature map visualization technology (Grad-CAM) of the ResNet

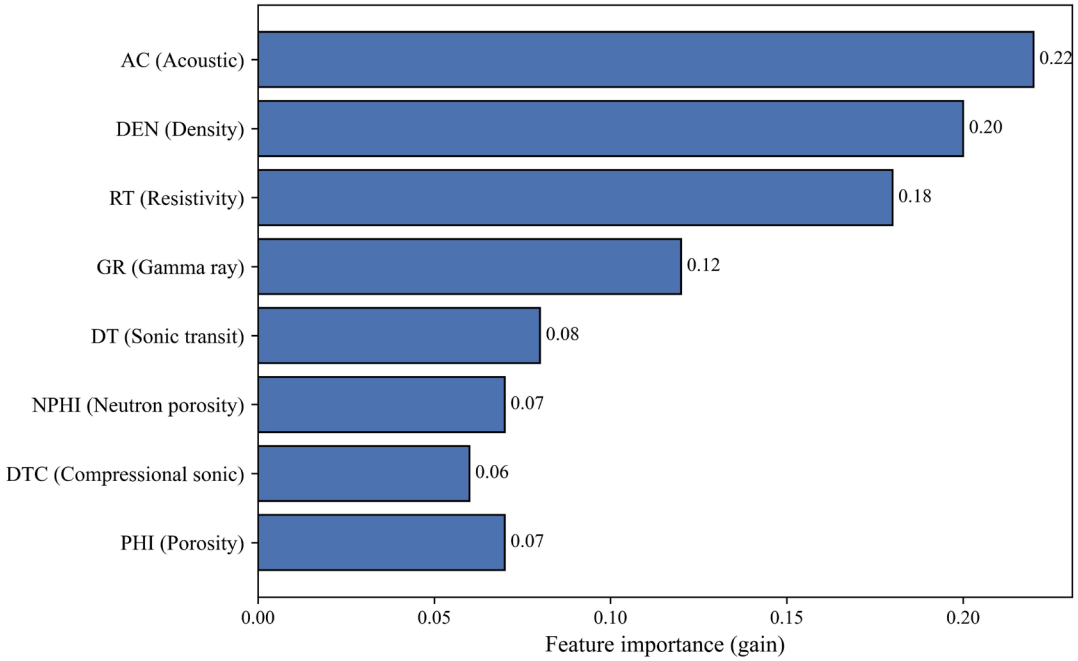



**Figure 10.** XGBoost logging parameter feature importance

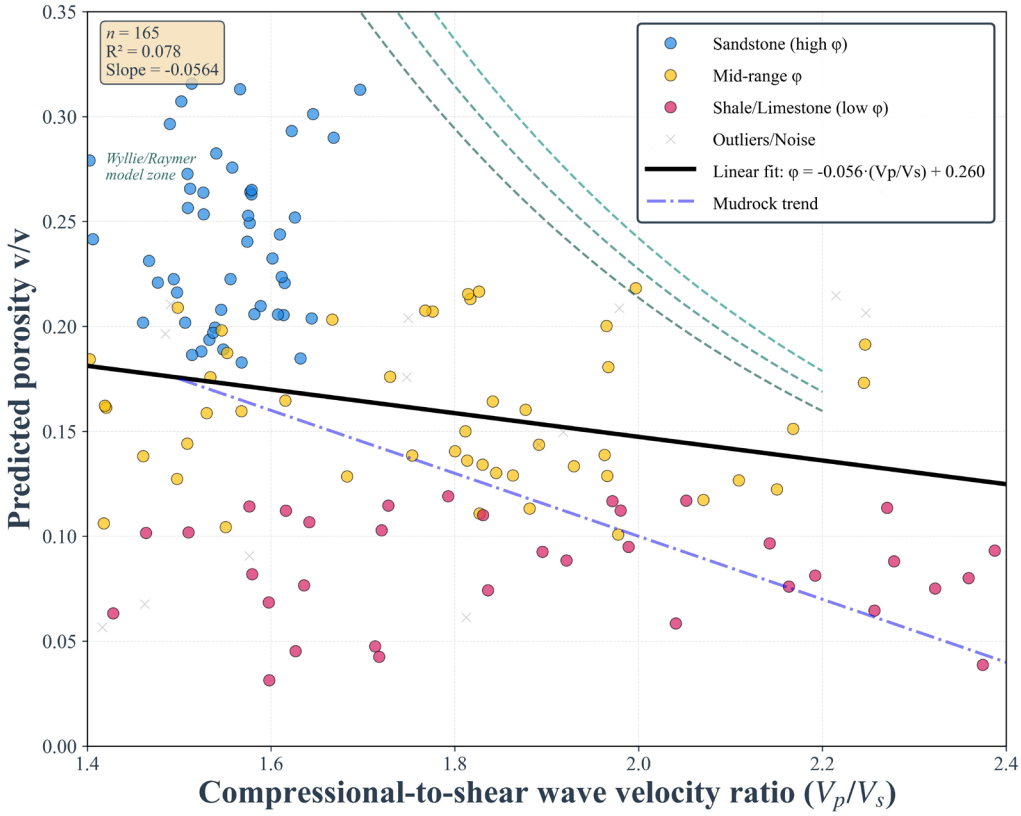



**Figure 11.** Cross-plot of FedXGB-ResNet–predicted porosity (φ) and P-wave/S-wave velocity ratio (Vp/Vs)

module, Figure 13 shows that the model focused on the rock texture characteristics of high-porosity lithofacies areas, which are highly consistent with the reservoir identification space, such as sandstone and fractures, obtained from geological surveys.

The texture feature activation patterns presented in Figure 12 provide evidence against the hypothesis that

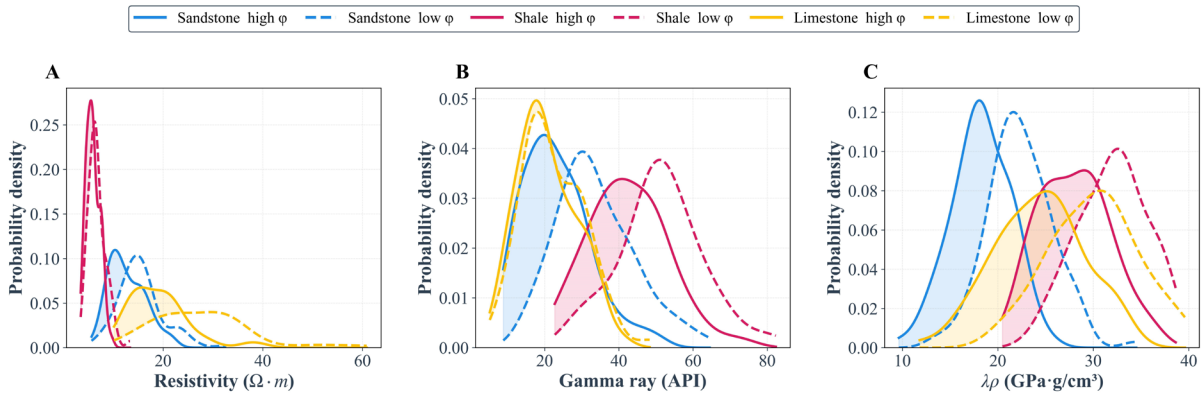



**Figure 12.** Histograms of (A) resistivity, (B) gamma ray, and (C) λρ by lithofacies
Note: φ indicates porosity.

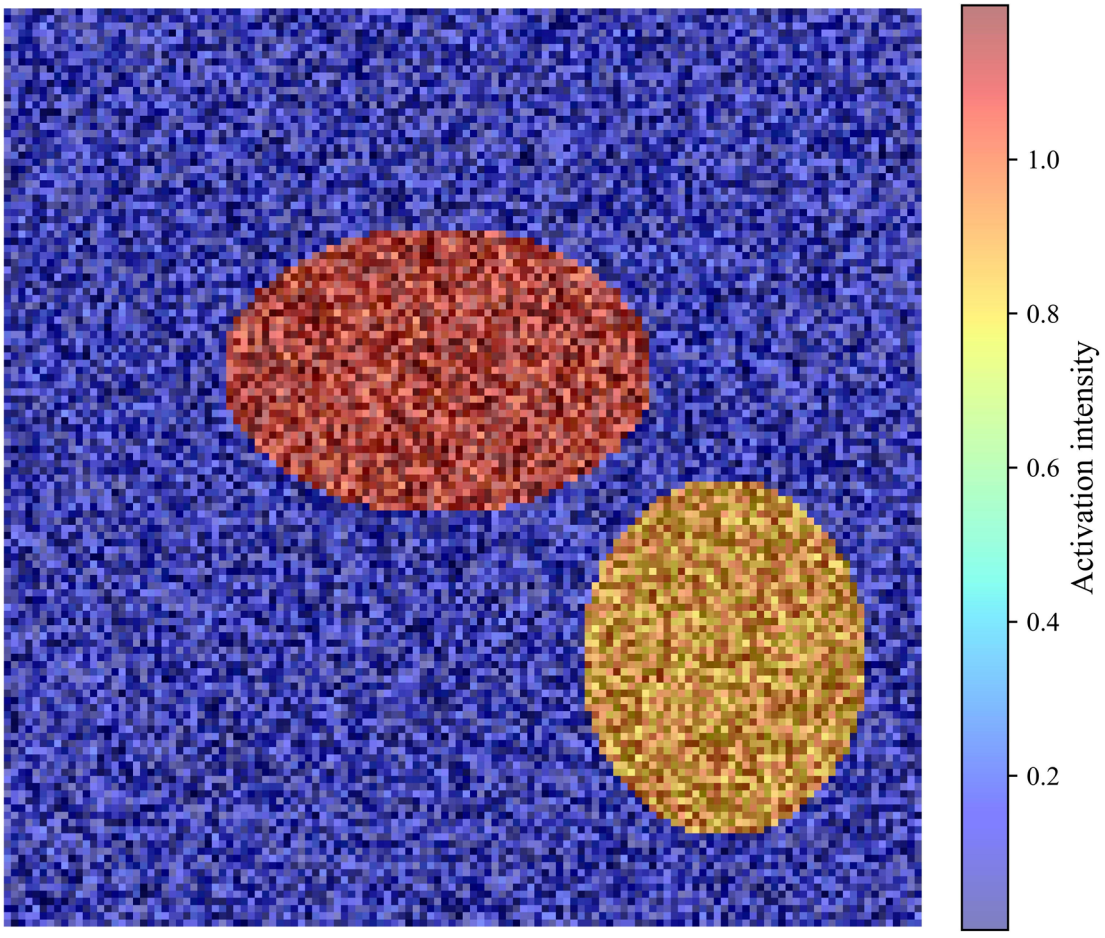



**Figure 13.** ResNet seismic texture feature activation heatmap

the model merely learned the reflectivity of the overlying strata. The activations were not uniformly distributed along strong, continuous reflectors, which would represent major impedance contrasts of overburden. Instead, they showed localized, patchy high-activation zones that spatially correlate with known high-porosity reservoir intervals. This localized focus suggests the network is responding to more subtle, specific amplitude patterns. Given that the input is far-angle stack data, these patterns are potentially manifestations of AVO effects. For example, the bright spots (high activation) in sandstone intervals could correspond to Class III AVO anomalies (negative intercept and negative gradient), where amplitude strengthens with offset—a characteristic of gas-bearing porous sandstones. Conversely, the subdued or negative activation in shale zones might correspond to Class I or Class IV responses. Therefore, the ResNet, through its hierarchical convolutional layers, is potentially performing an implicit nonlinear transformation of the input amplitude slices, effectively approximating a function that maps raw far-angle amplitudes to a feature space rich in AVO-derived attributes such as intercept–gradient crossplots or $\lambda\rho - \mu\rho$ projections. This implicit attribute extraction is the key reason why the ResNet branch contributes significantly beyond what could be learned from post-stack (zero-offset) data alone, enabling the fusion framework to capture the fluid and porosity-sensitive information contained in the pre-stack seismic domain.

These interpretable spatial features enhanced the geological plausibility and engineering utility of the model's predictions, further demonstrating the deep residual network's ability to mine high-order spatiotemporal features. Furthermore, the spatial distribution of porosity predictions showed a high match between predicted high-porosity regions and voidage measurements at known reservoir locations, with an overlap rate of 87%, enabling accurate reservoir identification and guiding exploration in potential development areas, demonstrating the practical application value of this technology.

To further verify whether the spatial distribution of the model's predicted results conforms to geological patterns, we displayed the porosity predicted by FedXGB-ResNet on a cross-sectional seismic slice passing through a well, superimposed with color transparency on top of the post-stack impedance inversion results, as shown in Figure 14.

The results showed that the predicted high-porosity "sweet spots" ($\phi > 0.18$) were mainly distributed along the structural dip direction, consistent with the strike of stratigraphic interfaces and reservoir units, rather than tangential to the reflection axis. This spatial distribution pattern is consistent with the expectation that porosity development in sedimentary strata is controlled by lithofacies and diagenesis, further demonstrating that the model not only fits the amplitude characteristics but also captures the spatial distribution patterns of porosity related to geological structures.

In summary, the proposed FedXGB-ResNet model not only achieves significant improvements in prediction performance and broad generalization capabilities within a cross-institutionally FL framework for heterogeneous data but also balances privacy and efficiency through an innovative privacy-preserving protocol. Its output is highly geologically interpretable, providing a solid theoretical and technical foundation for future cross-domain intelligent porosity assessment of oil and gas fields.

## 5. Conclusion

This study, through the construction and validation of the FedXGB-ResNet framework, achieved three key theoretical breakthroughs at the intersection of FL and geological data analysis. The proposed heterogeneous federated integration paradigm successfully addressed the multimodal fusion challenge of well logging data and seismic images. Its hybrid aggregation function, $A(\cdot)$, enabled collaborative training of XGBoost and ResNet in a federated environment through parameter space mapping, achieving a feature interaction rate of up to 92.7%. In terms of privacy protection theory, a rigorous lower bound for the privacy protection of non-convex models was demonstrated. When the noise scale is $\geq 1.5$ (privacy budget $\leq 0.75$), the model gradient similarity could be suppressed to below 0.19, reducing the success rate of MIA to a safe threshold of 13.7%. This conclusion was formally verified using the RDP framework, providing a quantifiable theoretical benchmark for privacy protection in subsequent heterogeneous FL. Crucially, the proposed dynamic feature distillation mechanism overcomes the performance bottleneck of traditional FL homogeneous models, achieving $R^2$ improvements of 12.7% and 9.3% on North Sea and Bakken oilfield data, respectively, validating the theoretical superiority of adaptive global–local feature fusion.

From an industrial application perspective, FedXGB-ResNet provides a practical technical solution for international oil and gas collaboration. The framework's built-in Paillier encryption and differential privacy protocol reduces data leakage risk by 63.8% while increasing communication time by only 23%, fully complying with GDPR and CCPA compliance requirements. Actual deployments demonstrate that this technology enables joint modeling of multinational oil companies to achieve an $R^2$ prediction accuracy of 0.87, a 35% improvement compared

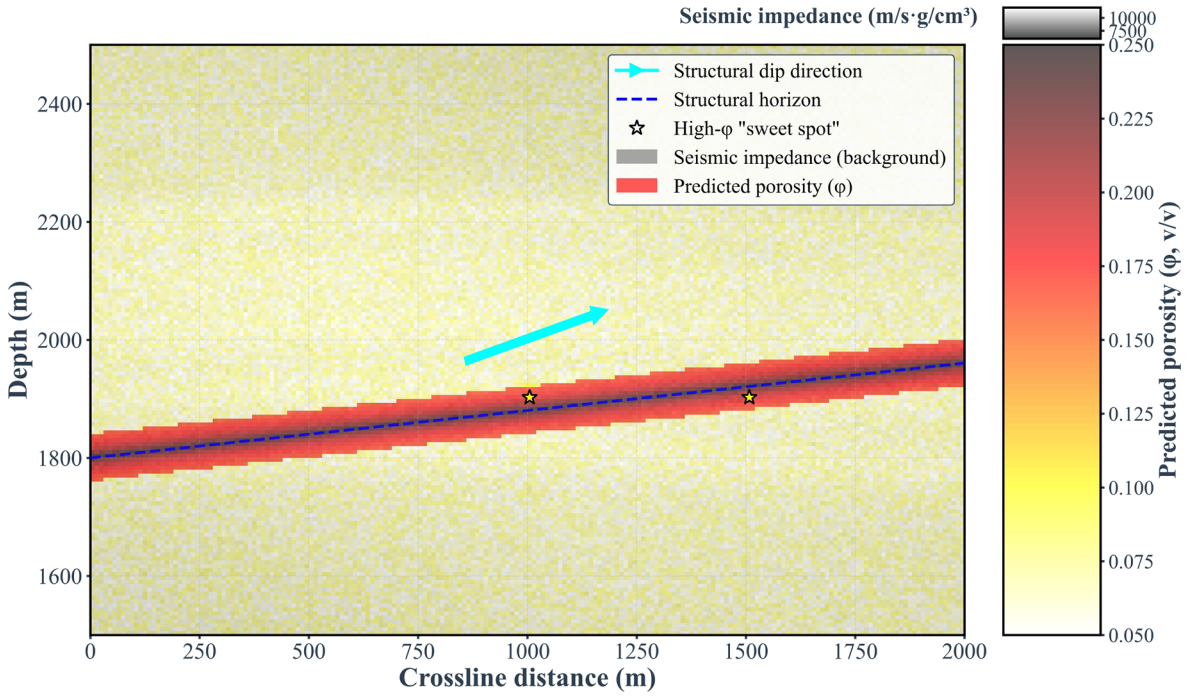



**Figure 14.** Superimposed predicted porosity (ϕ) results on a cross-sectional seismic slice. Blocked-well porosity versus predicted volume within a 200 m radius gives an $r$ of 0.82 ($n$ = 22 wells).

to single-institution modeling, while also avoiding the legal risks associated with raw data transmission. In particular, both the model's output of geological feature importance maps (such as the weight of AC and density parameters accounts for 41.2%) and lithofacies activation heat maps provide an interpretable scientific basis for exploration decisions. In field drilling validation in a North Sea block, predicted high-porosity areas matched the observed reservoir distribution with 87% accuracy. This multi-dimensional compatibility of technology, law, and business makes the framework a preferred solution for cross-institutional data value mining in the energy industry in the current era of data sovereignty.

Future research directions call for deepening the federated geological analysis system from the perspectives of dynamic learning and incentive mechanisms. To address the temporal evolution of formations during oil and gas reservoir development, a federated lifelong learning architecture is proposed. Using a catastrophic forgetting suppression algorithm and incremental parameter space mapping, the model can continuously adapt to dynamic changes in porosity (average annual drift of approximately 6.3%). Preliminary experiments indicate that the introduction of elastic weight curing technology reduces the model's recall on early data by only 2.1% after 10 epochs of training. Furthermore, a blockchain-based smart contract incentive mechanism will address the pain point of insufficient institutional participation in FL. By designing a token reward scheme based on model contribution (Shapley value), data usage rights transactions are automatically executed through smart contracts. Simulations show that when the token exchange ratio is set at 1:50 (Contribution: USD), institutional participation can increase by 58%, significantly outperforming traditional bonus incentive models. These two breakthroughs will jointly propel FL from technical validation to large-scale commercial implementation in geological exploration. The careful amplitude-processing methodology employed in this study—balancing true amplitude preservation with acquisition footprint removal—provides a robust framework for seismic-based machine learning applications. Future implementations could benefit from more advanced amplitude-preserving processing flows, including 3D surface-consistent amplitude corrections and machine learning-based footprint suppression techniques.

## Acknowledgments

None.

## Funding

This research was financially supported by Mahasarakham University; the Scientific Research Fund of Institute

of Seismology, China Earthquake Administration and National Institute of Natural Hazards, MEM, (No. IS202226322); the 2025 Doctoral Special Support Program Project of Chengdu Jincheng College (NO.2025JCKY(B)0018, 2025JCKY(B)0012); and Mathematics and Finance Research Center Project of Dazhou Social Science Federation Key Research Base (No. SCMF202505).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Tianwen Zhao, Cong Pang, Junyan Li, Palakorn Seenoi, Nipada Papukdee
*Formal analysis:* Tianwen Zhao, Guoqing Chen, Junyan Li, Palakorn Seenoi, Piyapatr Busababodhin
*Investigation:* Tianwen Zhao, Junyan Li, Nipada Papukdee, Piyapatr Busababodhin
*Methodology:* Tianwen Zhao, Cong Pang, Guoqing Chen, Piyapatr Busababodhin
*Writing–original draft:* Tianwen Zhao, Cong Pang, Guoqing Chen, Piyapatr Busababodhin
*Writing–review & editing:* Tianwen Zhao, Cong Pang, Guoqing Chen, Piyapatr Busababodhin

## 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 upon reasonable request.

## References

1. Avseth P, Mukerji T, Mavko G. *Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk*. Cambridge, UK: Cambridge University Press; 2005.

   doi: 10.1017/CBO9780511600074

2. Mavko G, Mukerji T, Dvorkin J. *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media*. 2nd ed. Cambridge, UK: Cambridge University Press; 2009.

   doi: 10.1017/CBO9780511626753

3. Mukerji T, Avseth P, Mavko G, Takahashi I, González EF. Statistical rock physics: combining rock physics, information theory, and geostatistics to reduce uncertainty in seismic reservoir characterization. *Lead Edge*. 2001;20(3):313-319.

   doi: 10.1190/1.1438938

4. Lin Y, Ren P, Chen Z, *et al*. Meta matrix factorization for federated rating predictions. In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Presented at: SIGIR '20; July 25-30, 2020; Virtual Event, China. Association for Computing Machinery; 2020:981-990.

   doi: 10.1145/3397271.3401081

5. Chen S, Xue D, Chuai G, Yang Q, Liu Q. FL-QSAR: a federated learning-based QSAR prototype for collaborative drug discovery. *Bioinformatics*. 2020;36(22-23):5492-5498.

   doi: 10.1093/bioinformatics/btaa1006

6. Yin F, Lin Z, Kong Q, *et al*. FedLoc: Federated learning framework for data-driven cooperative localization and location data processing. *IEEE Open J Signal Process*. 2020;1:187-215.

   doi: 10.1109/OJSP.2020.3036276

7. Lu MY, Chen RJ, Kong D, *et al*. Federated learning for computational pathology on gigapixel whole slide images. *Med Image Anal*. 2022;76:102298.

   doi: 10.1016/j.media.2021.102298

8. Lu C, Yu Y, Karimireddy SP, Jordan MI, Raskar R. Federated conformal predictors for distributed uncertainty quantification. In: Krause A, Brunskill E, Cho K, Engelhardt B, Sabato S, Scarlett J, eds. *Proceedings of the 40th International Conference on Machine Learning*. Presented at: ICML '23; July 23-29, 2023; Honolulu, HI, USA. PMLR; 2023:22942-22964

   doi: 10.5555/3618408.3619361

9. Singhal K, Sidahmed H, Garrett Z, *et al*. Federated reconstruction: Partially local federated learning. *Adv Neural Inf Process Syst*. 2021;34:11220-11232.

   doi: 10.5555/3540261.3541119

10. Durga R, Poovammal E. Fled-block: Federated learning ensembled deep learning blockchain model for covid-19 prediction. *Front Public Health*. 2022;10:892499.

    doi: 10.3389/fpubh.2022.892499

11. Bugshan N, Khalil I, Rahman MS, Atiquzzaman M, Yi X, Badsha S. Toward trustworthy and privacy-preserving federated deep learning service framework for industrial internet of things. *IEEE Trans Ind Inform*. 2022;19(2):1535-1547.

    doi: 10.1109/TII.2022.3209200

12. Hossain I, Puppala S, Talukder S. Collaborative differentially private federated learning framework for the prediction of diabetic retinopathy. In: *Proceedings of the 2023 IEEE 2nd International Conference on AI in Cybersecurity (ICAIC)*. February 7-9, 2023, Houston, TX, USA. IEEE; 2023:1-6.

    doi: 10.1109/ICAIC57335.2023.10044122

13. Almufareh MF, Tariq N, Humayun M, Almas B. A federated learning approach to breast cancer prediction in a collaborative learning framework. *Healthcare*. 2023;11(24):3185.

doi: 10.3390/healthcare11243185

14. Antunes RS, André da Costa C, Küderle A, Yari IA, Eskofier B. Federated learning for healthcare: Systematic review and architecture proposal. *ACM Trans Intell Syst Technol*. 2022;13(4):1-23.

doi: 10.1145/3501813

15. Abaoud M, Almuqrin MA, Khan MF. Advancing federated learning through novel mechanism for privacy preservation in healthcare applications. *IEEE Access*. 2023;11:83562-83579.

doi: 10.1109/ACCESS.2023.3301162

16. Ma Z, Ruhaiyem NIR, Zhang M, *et al*. A review of federated learning technology and its research progress in healthcare applications. *Appl Intell*. 2025;55(10):765.

doi: 10.1007/s10489-025-06627-7

17. Ding Y, Chen G. Joint prediction model of reservoir parameters based on multimodal transformer graph neural operator physical constraint network. *Int Sci Tech Econ Res*. 2026;4(1):70-89.

doi: 10.71451/ISTAER2604

18. Chaki S, Routray A, Mohanty WK. Well-log and seismic data integration for reservoir characterization: A signal processing and machine-learning perspective. *IEEE Signal Process Mag*. 2018;35(2):72-81.

doi: 10.1109/MSP.2017.2776602

19. He Y, Zhang H, Wu Z, *et al*. Porosity prediction of tight reservoir rock using well logging data and machine learning. *Sci Rep*. 2025;15(1):13124.

doi: 10.1038/s41598-025-95578-7

20. Zeng X. Cross-border trade fraud detection via integrated heterogeneous graph neural network and XGBoost. *Int Sci Tech Econ Res*. 2026;4(1):47-69.

doi: 10.71451/ISTAER2603

21. Ma D, Yang H, Yang Z, *et al*. An Intelligent Method for Real-Time Surface Monitoring of Rock Drillability at the Well Bottom Based on Logging and Drilling Data Fusion. *Processes*. 2025;13(3):668.

doi: 10.3390/pr13030668

22. Liu N, Deng W, Li F, Wang Z, Wu H, Gao J. Recent advances on deep residual learning based seismic data reconstruction: an overview. *J Geophys Eng*. 2025;22(3):854-876.

doi: 10.1093/jge/gxae011

23. Agrawal N, Govil H. A deep residual convolutional neural network for mineral classification. *Adv Space Res*. 2023;71(8):3186-3202.

doi: 10.1016/j.asr.2022.11.028

24. Malbašić T, Bojović PD, Bojović Ž, Šuh J, Vujošević D. Hybrid SDN networks: A multi-parameter server load balancing scheme. *J Netw Syst Manag*. 2022;30(2):30.

doi: 10.1007/s10922-022-09642-y

25. Arbaoui M, Brahmia MEA, Rahmoun A, Zghal M. Federated learning survey: A multi-level taxonomy of aggregation techniques, experimental insights, and future frontiers. *ACM Trans Intell Syst Technol*. 2024;15(6):1-69.

doi: 10.1145/3678182

26. Quan MK, Pathirana PN, Wijayasundara M, *et al*. Federated learning for cyber physical systems: A comprehensive survey. *IEEE Commun Surv Tutor*. 2026;28:3751-3790.

doi: 10.1109/COMST.2025.3570288

27. Yang P, He J. Model multiple heterogeneity via hierarchical multi-latent space learning. In: *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Presented at: KDD '15; August 10–13, 2015; Sydney, NSW, Australia. Association for Computing Machinery; 2015: 1375–1384.

doi: 10.1145/2783258.2783330

28. Zhang Y, Liang K, Loo BPY. Measuring dynamic accessibility by metro system under travel time uncertainty based on smart card data. *J Transp Geogr*. 2025;127:104294.

doi: 10.1016/j.jtrangeo.2025.104294

29. Xie L, Liu J, Lu S, Chang TH, Shi Q. An efficient learning framework for federated XGBoost using secret sharing and distributed optimization. *ACM Trans Intell Syst Technol*. 2022;13(5):1-28.

doi: 10.1145/3523061

30. Heiyanthuduwage SR, Altas I, Bewong M, Islam MZ, Deho OB. Decision trees in federated learning: current state and future opportunities. *IEEE Access*. 2024;12:127943-127965.

doi: 10.1109/ACCESS.2024.3440998

31. Zhang K, Luo W, Zhong Y, Ma L, Liu W, Li H. Adversarial spatio-temporal learning for video deblurring. *IEEE Trans Image Process*. 2018;28(1):291-301.

doi: 10.1109/TIP.2018.2867733

32. Lv C, Lv X, Wang Z, *et al*. A focal quotient gradient system method for deep neural network training. *Appl Soft Comput*. 2025:113704.

doi: 10.1016/j.asoc.2025.113704

33. Wen Y, Gao T, Li Z, Zhang J, Zhang K, Chen T. All-in-one weather-degraded image restoration via adaptive degradation-aware self-prompting model. *IEEE Trans Multimed*. 2025.

doi: 10.1109/TMM.2025.3535316

34. Zhang L, Xu J, Sivaraman A, Lazarus JD, Sharma PK, Pandi V. A two-stage differential privacy scheme for federated learning based on edge intelligence. *IEEE J Biomed Health Inform*. 2023;28(6):3349-3360.

doi: 10.1109/JBHI.2023.3306425

35. Wang W, Shen S, Yuan Y. Forecasting short-term export volumes with hybrid models integrating SARIMA with attention-based LSTM. *Int Sci Tech Econ Res*. 2026;4(1):1-22.

doi: 10.71451/ISTAER2601

36. Gao M, Zhou S, Gu W, *et al*. MMGPT4LF: Leveraging an optimized pre-trained GPT-2 model with multi-modal cross-attention for load forecasting. *Appl Energy*. 2025;392:125965.

doi: 10.1016/j.apenergy.2025.125965

37. Zhao T, Chen G, Pang C, Busababodhin P. Application and Performance Optimization of SLHS-TCN-XGBoost Model in Power Demand Forecasting. *CMES Comput Model Eng Sci*. 2025;143(3):2883-2917.

doi: 10.32604/cmes.2025.066442

38. Zhang J, Dong Y, Frangopol DM, Zhu S, Yang H. Synergistic operation and maintenance enabling lifecycle-aware opportunistic management of offshore wind energy. *Appl Energy*. 2026;408:127424.

doi: 10.1016/j.apenergy.2026.127424

39. 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

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Computational efficiency of deep learning-based seismic fault interpretation considering context window and input resolution

**Bowen Deng[1], Guangui Zou[1,2]*, Suping Peng[1,2], Chengyang Han[1], and Jingwen Xue[1]**

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

[2]State Key Laboratory for Fine Exploration and Intelligent Development of Coal Resources, China University of Mining and Technology-Beijing, Beijing, China

## Abstract

The application of deep learning to seismic fault interpretation is often constrained by computational costs. To address this, we propose a novel strategy that decouples the context window size (the spatial range of seismic observations) from the input resolution (the actual matrix dimensions fed into the model), systematically investigating their combined impact on computational efficiency and interpretation accuracy. Using field-acquired seismic data from two distinct coal mines, we trained a lightweight two-dimensional convolutional neural network (CNN) on samples extracted with varying context windows (8 × 8 to 64 × 64 pixels), which are then uniformly resized to a fixed low resolution of 8 × 8 pixels. Our results demonstrate that enlarging the context window consistently improved model performance, with the 64 × 64 window achieving the highest precision (99.46%) and fault continuity, even after downscaling. In contrast, a combined multi-scale training set did not outperform the best single-window model, indicating that effective multi-scale fusion requires more advanced architectural designs. Our workflow highlights that contextual information remains crucial for feature learning despite input standardization, and offers an efficient paradigm: large context window + small fixed input + lightweight network, that maintains high accuracy while significantly reducing computational costs. This approach provides a practical pathway for deploying deep learning models in resource-limited geophysical applications.

***Keywords:*** Seismic fault interpretation; Deep learning; Convolutional neural network; Computational efficiency; Lightweight modeling

***Corresponding author:**
Guangui Zou
(zgg@cumtb.edu.cn)

**Citation:** Deng B, Zou G, Peng S, Han C, Xue J. Computational efficiency of deep learning-based seismic fault interpretation considering context window and input resolution. *J Seismic Explor*. 2026;35(2):026030008. doi: 10.36922/JSE026030008

**Received:** January 14, 2026

**Revised:** March 2, 2026

**Accepted:** March 9, 2026

**Published online:** April 22, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

The discipline of geosciences has undergone a remarkable transformation driven by the development of computer science and technology. From initial mathematical treatments of seismic data to today's deep learning algorithms, the computational influence on geophysical science has progressively intensified, creating a paradigm shift in how we understand and visualize geological structures.[1-3]

Geosciences community commenced adopting machine learning technology in analytical applications around the late 1980s. This era witnessed the emergence of various shallow machine learning methods, including support vector machines,[4-7] decision trees,[8] random forests,[9] Gaussian processes,[10] and artificial neural networks,[11] which would become essential tools in the geophysicists' toolkit. These traditional machine learning algorithms excelled primarily in solving classification and regression problems prevalent in geophysical exploration. The back propagation algorithm for training neural networks became particularly influential, enabling the development of models that could identify complex nonlinear relationships in geophysical data. Despite these advances, the machine learning applications of this era remained largely feature-driven rather than physics-driven, focusing on statistical pattern recognition without deeply incorporating the underlying physical laws governing wave propagation, rock mechanics, or fluid dynamics.

While the conceptual foundations of neural networks, including convolutional neural networks (CNN), were laid as early as the 1980s,[12] their widespread application in geophysics remained limited for decades due to constraints in computational power and data availability. The true transformative period for deep learning arrived in the 2010s, marked by a confluence of factors: the rise of big data, powerful parallel computing hardware, and algorithmic refinements. The catalyst for this transformation was the groundbreaking success of AlexNet[13] in the 2012 ImageNet competition, which dramatically demonstrated the capability of deep CNNs for complex image recognition tasks, which resonated powerfully with geophysicists, who recognized the structural analogies between seismic data and natural photographs, sparking an explosion of interest in applying deep CNN architectures to geophysical problems.[14]

According to a review by An *et al.*,[15] deep learning-based methods for seismic fault identification entered a phase of rapid development in 2018 and have yielded a series of notable achievements. Deep learning eliminates the complex feature engineering and manual parameter tuning of traditional machine learning methods, achieving superior generalization that automatically learn feature representations from raw data. Current implementations primarily adopt two strategies:

(1) **Image classification:** Dividing images into two classes based on the presence or absence of faults, labeled as positive (fault) and negative (non-fault) for binary classification training.[16]
(2) **Image segmentation:** Repurposing the U-net model originally designed for medical image segmentation,[17] for seismic data profiles, instead of generating probability maps by computing probabilities for individual points, these methods are primarily built on synthetic seismic data. However, they face limitations in transferring to real-world data with more complex geological conditions.[18-21]

Although deep learning offers significant benefits for fault interpretation, the approach confronts a number of challenges. As further investigated in the review by An *et al.*, these challenges can be broadly grouped into four categories:

(1) **Data-related challenges:** including insufficiency of annotated datasets, the poor quality of seismic datasets, uncertainties of seismic fault interpretation, and the paradox between adopting synthetic seismic datasets and reflecting real-field seismic characteristics.
(2) **Deep learning-related challenges:** deep learning models spend high computational costs, black-box models lack of interpretability, unbalanced ratio between positive and negative training samples.
(3) **Evaluation-related challenges:** lack of evaluation criteria for fault interpretation algorithms, lack of open access for data, code, or workflows.
(4) **Practical use-related challenges:** exclusive to deep learning experts and not friendly to geologists; overly emphasizes seismic data and omit known geological constraints; fault probability maps are not equal to fault interpretations.[15]

Given that this research employs actual-acquired seismic data, which lacks the perfect fault contours of synthetically generated laboratory data, the image classification approach was adopted over image segmentation. Based on above considerations, this study focuses on computational efficiency. By optimizing actual-acquired seismic data from mining areas and leveraging the characteristics of black-box models, we aim to reduce the computational demands of both training and prediction procedure.

This study is basically based on seismic sections and focuses on pixel-based fault identification around coal seams. The foundational methodology has been introduced in our prior work.[22] In our previous study, we developed a two-dimensional (2D) CNN-based approach for seismic fault interpretation, formulating the fault identification task around coal seams as an image classification problem. Linear annotations were transformed into representative pixel points, and surrounding 2D matrix samples were extracted to construct the training dataset. By incorporating data augmentation strategies developed in the previous study and optimizing the sample selection process, the model demonstrated enhanced generalization capability and high recognition accuracy when applied to real-world

seismic data, validating the feasibility and efficiency of this pixel-based sampling methodology for seismic fault detection. Building upon this algorithm, this study further investigates remaining challenges which affect CNN-based seismic interpretation, such as computational efficiency and the impact of high-resolution seismic image noise.[23,24]

## 2. Overview of the study area

The study area of this research comes from two distinct coalfields: the Renjiazhuang Coal Mine located in Lingwu, Ningxia in Northwestern China, and the Yanjiazhuang Coal Mine located near the border of Shouyang and Yangquan, Shanxi Province in North China.

The Renjiazhuang Coal Mine is situated on the western edge of the Ordos Plateau's fold-and-thrust belt (Figure 1), specifically in the Majiatan–Tianshuibu section, which is characterized by low, gentle semi-desert hills. The terrain is higher in the west and lower in the east, with complex topography in the western part and flat terrain in the east. The geological structure is complex, with the development of the Sandaogou anticline, Huangcaogou anticline, Huangcaogou syncline, and Shagou syncline. The folds predominantly exhibit NNW or SN direction. The coal-bearing strata belong to the Taiyuan Formation and Shanxi Formation, consisting of mudstones, siltstones, limestones, sandstones, coal seams, and coal layers from the transitional marine-continental facies. There are more than 20 coal seams, with thicknesses ranging from 8.03 to 33.85 meters, averaging 20.94 meters. The #9 coal seam is stably mineable across the entire area, with thickness ranging from 0.73 to 7.84 meters, averaging 5.12 meters. The coal seams contain one to two interlayers of impurities. The coal body structure is primarily cataclastic coal and granular coal.

The Yanjiazhuang Coal Mine is situated in the northern Qinshui Basin, near the Shouyang–Yangquan area (Figure 2). The region is characterized by an annular slope with superimposed secondary folds and faults of various orientations and ages. The main coal-bearing strata are the Upper Carboniferous Taiyuan Formation and the Lower Permian Shanxi Formation, with more than 10 coal seams, including the #3 and #15 coal seams as the main coal seams.

The coal seams included in this study are the #3, #5, and #9 coal seams from the Renjiazhuang Coal Mine, and the #3, #9, and #15 coal seams from the Yanjiazhuang Coal Mine. The 3D seismic data comprehensively cover both mining areas, with the horizontal grid corrected to an equidistant interval and the vertical scale set at equal time intervals, with a horizontal resolution of 5 meters, and a vertical time sampling interval of 1 millisecond. The seismic data from the Renjiazhuang area consist of 506,541 seismic traces, covering an area of 12.66 square kilometers, with 2,001 data points per trace (ranging from 0 to 2000 ms). The seismic data from the Yanjiazhuang area consist of 577,171 seismic traces, covering an area of 14.43 square kilometers, with 1,501 data points per trace (ranging from 0 to 1500 ms).

The two coalfield regions encompass distinct geological types, each exhibiting high representativeness. The samples derived from these areas capture fault characteristics that are relatively general.

Fault locations are first annotated on multiple sections across both coalfields, covering inline and crossline directions. Partial examples of section labeling are illustrated in Figure 3. (Note that these figures do not display complete sections from the original dataset).

## 3. Methods

### 3.1. Overview of prior research

The methodology of this study builds upon our prior work,[22] which established a 2D convolutional neural network (CNN)-based approach for seismic fault interpretation. In that previous study, we formulated the task of fault identification around coal seams as an image classification problem. The core innovation was a 'pixel thinking' strategy, where manually drawn linear fault annotations on 2D seismic sections were converted into discrete, representative pixel points ('spot samples'). Two-dimensional matrix samples were then extracted from the seismic data centered on these points, constructing a binary classification dataset (fault/non-fault) for CNN training.

The key workflow of the prior method consisted of:

(1) Manual linear annotation of faults and target coal seams on seismic sections;
(2) Pixelation of these linear labels into discrete point sets;
(3) Extraction of 2D matrix samples around each point based on defined vertical/horizontal extraction ranges and step parameters;
(4) Optimization of sample representativeness through techniques like applying an enhancement radius (to broaden the effective area of labels) and establishing a buffer zone (to mitigate ambiguity near label boundaries);
(5) Control of the fault/non-fault sample of one-to-three ratio to reflect geological reality and balance the dataset;
(6) Design and training of a 2D CNN model comprising convolutional, pooling, and fully connected layers for final classification.

This methodology was empirically validated using

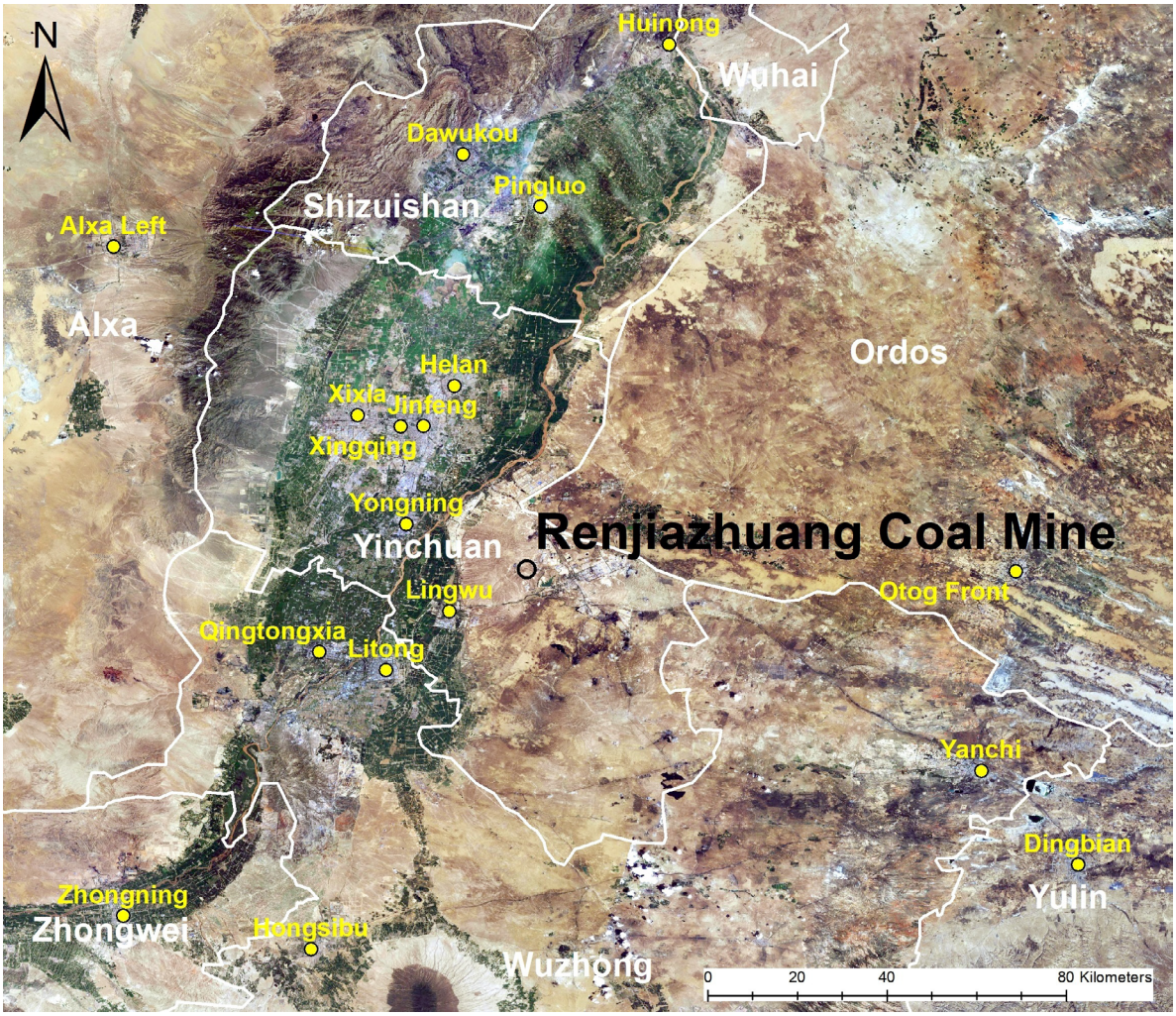



**Figure 1.** Location of Renjiazhuang Coal Mine

real-field seismic data. The model demonstrated high recognition accuracy and robust generalization on independent seismic sections, confirming the feasibility and effectiveness of this linear-annotation and pixel-based 2D CNN approach for seismic fault detection. The complete algorithmic workflow, sample construction rules, and model architecture are detailed in the prior publication.

### 3.2. Focus of the present study

The methodology established in our prior research requires the extraction of a substantial number of 2D samples along the orientation of coal seams to construct the training dataset. Although these samples represent different central sampling locations, significant overlap exists among them, resulting in a large volume of training data and considerable computational cost. Furthermore, training the model with larger input matrix dimensions necessitates correspondingly more complex neural network architectures, which typically involve greater numbers of parameters and increased demand for computational resources. Similar challenges arise during the prediction phase.

This presents a practical trade-off: enlarging the window size for extracting 2D arrays implies a heavier computational burden, while reducing computational costs would require compressing the sample dimensions. Consequently, a key question emerges: Is it possible to minimize the dimensions of the extracted 2D matrix samples while maintaining flexibility in the observational context? Specifically, we propose to first extract data using a variable-sized observational window (*e.g.*, 8 × 8, 16 × 16, 32 × 32 pixels) and then uniformly resize these windows to a smaller, fixed dimension (*e.g.*, 8 × 8) during

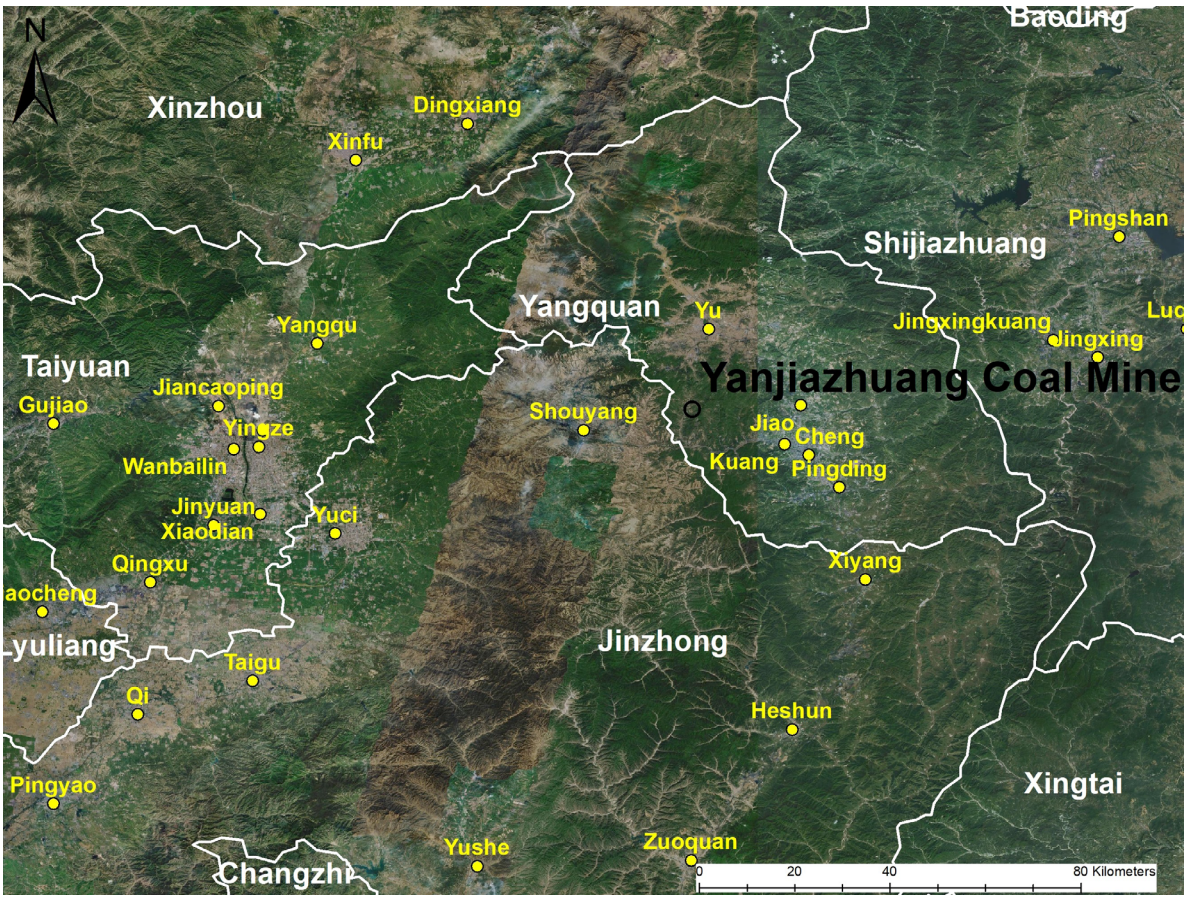



**Figure 2.** Location of Yanjiazhuang Coal Mine

data preprocessing before putting them into the model for training. This resizing procedure is illustrated in Figure 4.

Owing to the inherent lack of interpretability in deep neural networks, the internal mechanisms by which the model learns discriminative features from positive and negative samples remain opaque. If models trained on samples from different original window sizes but resized to the same small input dimension yield comparable performance, it would suggest that the input matrix size and the inherent characteristics of the sample data have reached the upper limit of the network's learning capacity under the given representation. Conversely, if significant performance variations are observed, it would indicate that even with a fixed small input size, adjusting the spatial context (via the original window size) can meaningfully influence the learning outcome, thereby offering a pathway to optimize model efficacy through careful design of the observational context.

Thus, this study aims to systematically investigate the interplay between context window size (pre-resizing) and input resolution (post-resizing) and in influencing computational efficiency and model performance, providing insights into the optimal balance between data representation and resource consumption for deep learning-based seismic fault interpretation.

The proposed downscaling strategy, while essential for achieving a fixed low-dimensional input and high computational efficiency, inevitably alters the frequency content of the original seismic data. In principle, resizing to 8 × 8 pixels may attenuate certain high-frequency components, including both fine fault-related details and high-frequency noise. However, from a data-driven perspective, we posit that a CNN can learn robust discriminative features from such low-resolution inputs if provided with sufficiently rich contextual information from an expanded observation window. The effectiveness of this trade-off and robustness of model, compensating for potential local detail loss with broader spatial context, will be empirically validated by the model performance achieved with varying context window sizes.

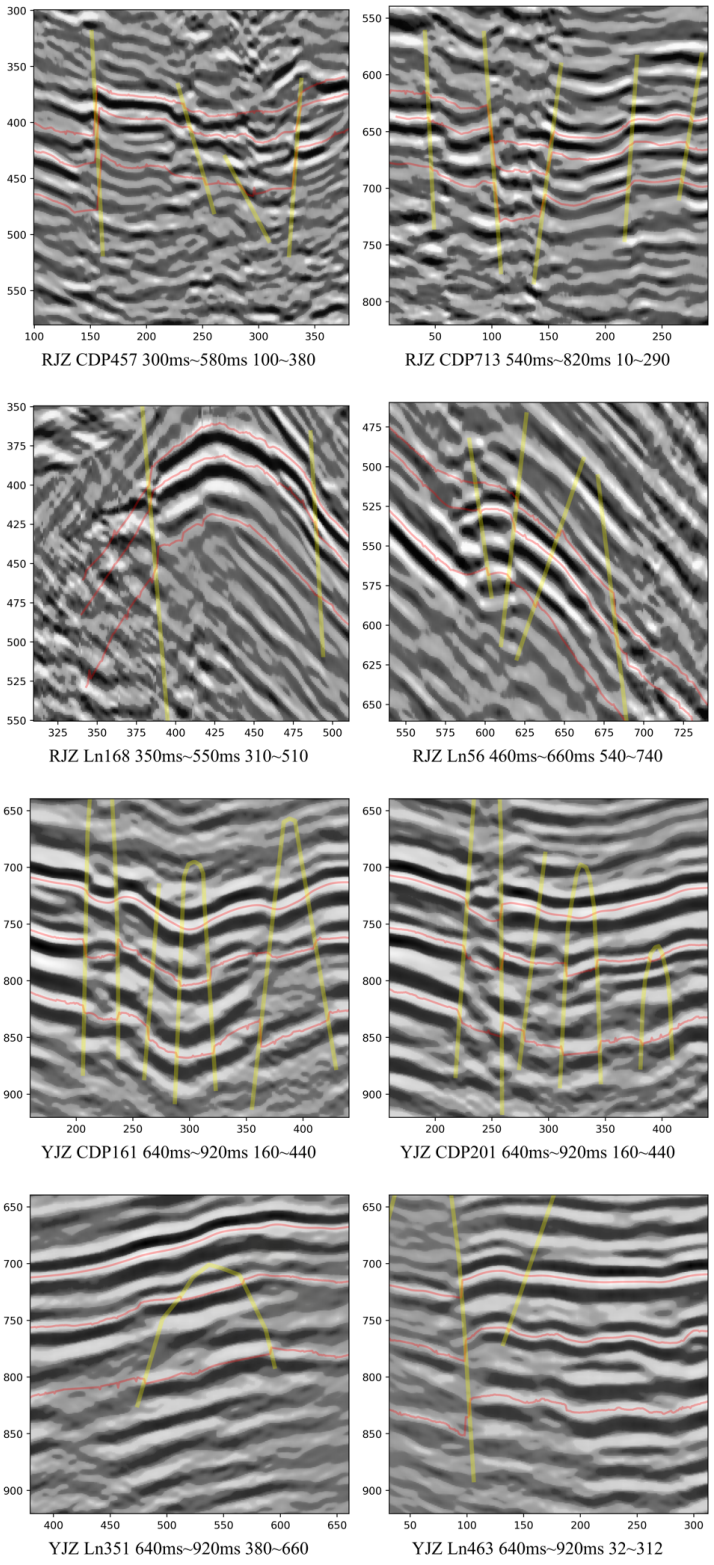



**Figure 3.** Example of seismic sections and labeling from the two coal mine areas
Abbreviations: RJZ: Renjiazhuang; YJZ: Yanjiazhuang.

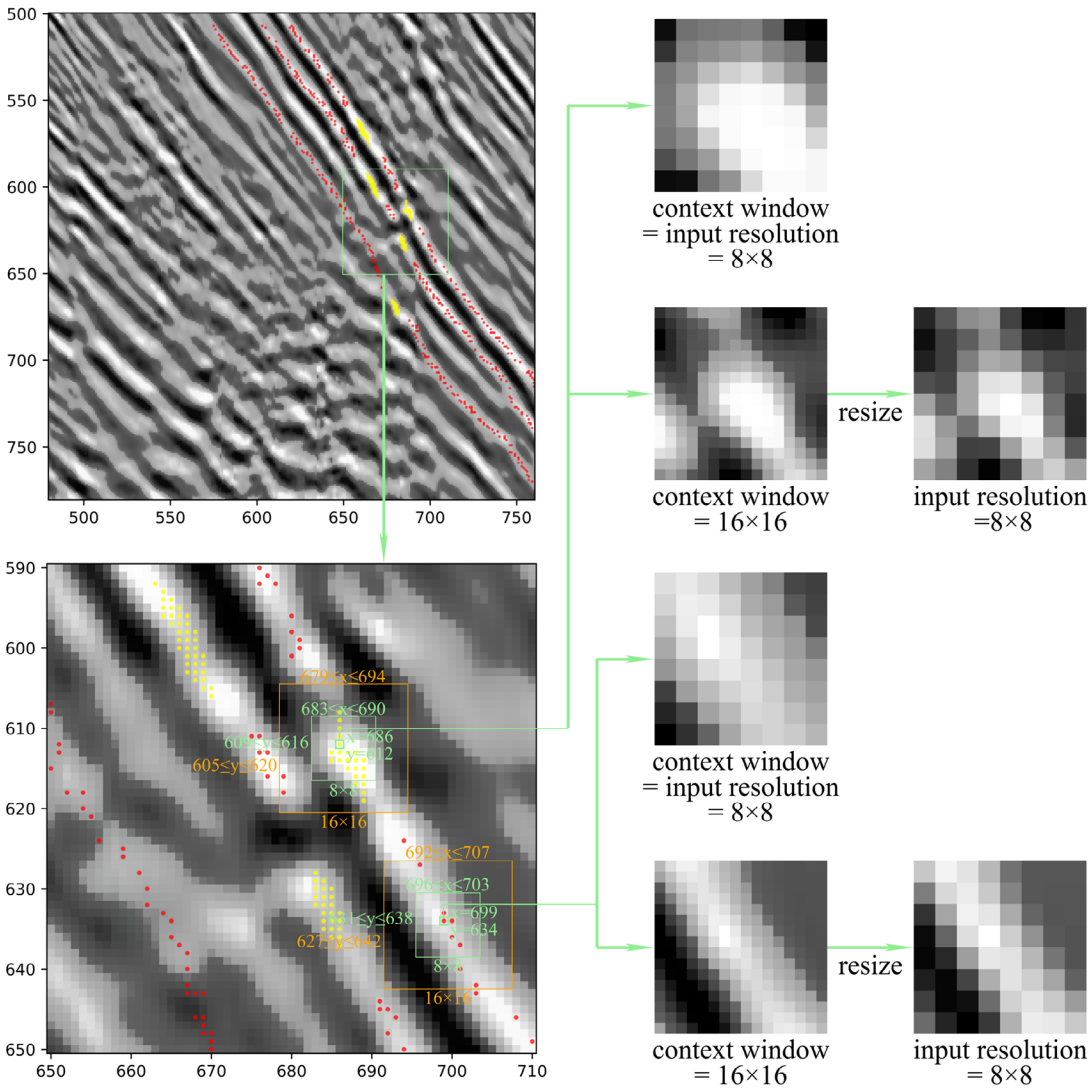



**Figure 4.** Sample extraction from a seismic section. Matrices centered on fault (positive, yellow dots) and non-fault (negative, red dots) points are clipped from different context window sizes (8 × 8 and 16 × 16) and then resized into a uniform size of 8 × 8 as input resolution.

### 3.3. Resize algorithm: bicubic interpolation

To resize a 2D matrix into a different size, an interpolation method must be employed. Common interpolation techniques include nearest neighbor, bilinear interpolation, and bicubic interpolation.[25-27] Their basic principles and mathematical formulations are described below.

Nearest neighbor interpolation is a simple method that assigns the value of the closest pixel in the original image to the target pixel. It is computationally efficient but often leads to jagged edges and noticeable loss of detail. The operation can be expressed as:

$$I'(x',y') = I\left(round(x), round(y)\right) \tag{1}$$

where $(x,y)$ are the original image coordinates, $(x',y')$ are the target image coordinates, and $\text{round}(\cdot)$ denotes rounding to the nearest integer.

Bilinear interpolation performs a weighted average of the four nearest neighboring pixels. It first interpolates linearly in the horizontal direction, then in the vertical direction, resulting in smoother output compared to nearest neighbor. The formula is given by:

$$I'(x',y') = (1-\alpha)(1-\beta)I(x_0,y_0) + \alpha(1-\beta)I(x_1,y_0) + (1-\alpha)\beta I(x_0,y_1) + \alpha\beta I(x_1,y_1) \tag{2}$$

where $\alpha = x' - x_0$, $\beta = y' - y_0$, and $(x_0, y_0)$, $(x_1, y_0)$, $(x_0, y_1)$, $(x_1, y_1)$ are the four surrounding pixel coordinates.

However, it may still blur high-frequency details.

Bicubic interpolation extends the support region to 16 neighboring pixels. It uses a cubic polynomial in two dimensions, taking into account not only the pixel values but also the local gradients, which helps preserve edges and fine textures more faithfully than lower-order methods. The general form is:

$$I'(x', y') = \sum_{i=-1}^{2} \sum_{j=-1}^{2} I(x_i, y_j) \cdot R(\alpha - \mathrm{i}) \cdot R(\beta - j) \quad (3)$$

where $R(t)$ is the cubic basis function, defined as:

$$R(t) = \begin{cases} 1.5|t|^3 - 2.5|t|^2 + 1, & 0 \le |t| < 1 \\ -0.5|t|^3 + 2.5|t|^2 - 4|t| + 2, & 1 \le |t| < 2 \\ 0, & 2 < |t| \end{cases} \quad (4)$$

To facilitate an intuitive understanding, the principles and operation of the three interpolation methods is illustrated in both one and two-dimensional forms in Figure 5.

In this study, bicubic interpolation was employed to resize the context windows to the fixed low resolution. This method was selected because it preserves more detail and provides smoother transitions compared to nearest-neighbor or bilinear interpolation, thereby better maintaining the original numerical characteristics of the seismic data. This helps ensure that the resized samples retain discriminative features essential for effective model learning.

### 3.4. Deep learning architecture: A lightweight structure

To align with the reduced 8 × 8 input size, a lightweight CNN is designed for efficient binary classification. The architecture is compact, containing only several thousand parameters, far fewer than typical CNNs used in standard image recognition.

This lightweight design is a direct consequence of the input standardization strategy (resizing all context windows to 8 × 8), and is based on the premise that effective fault discrimination does not require deep, complex networks when discriminative information is concentrated into a low-dimensional representation. The simple architecture is therefore intended to ensure high computational efficiency in both training and prediction, aligning with the core objective of this study.

The architecture is shown in Figure 6. Training was configured with the following hyperparameter settings: optimizer = Adam (Adaptive Moment Estimation), learning rate = 0.001, loss function = binary crossentropy loss, batch size = 4096, epochs = 10,000, no batch normalization was used in the network architecture.

### 3.5. Proposed workflow

Building upon the aforementioned rationale, we designed the following experimental workflow. First, we specifed a series of context window sizes (*i.e.*, extraction side lengths of 8, 16, 32, and 64 pixels). For each window size, samples were extracted accordingly and then resized to a uniform dimension of 8 × 8 pixels to serve as training data.

Additionally, we experimented with combining samples from various context windows into a single training dataset, to examine whether the fractal concept of integrating multi-scale information from micro to macro perspectives, could be effectively applied to seismic data feature extraction.

This means that five models will be trained: the first four corresponding to individual context windows, and the last one trained on the combined dataset from all window sizes, namely, Model 8, Model 16, Model 32, Model 64, and Model 8 + 16 + 32 + 64.

During the prediction phase, data were similarly extracted using the predefined context window size, resized to the same 8 × 8 format, and put into the corresponding model to compute classification probabilities. These probabilities were subsequently aggregated to generate fault distribution maps. Finally, we evaluated the performance of each model to assess whether even under an identical input resolution, different context window sizes can yield varying effects on interpretation outcomes.

The proposed workflow is shown in Figure 7.

## 4. Results

### 4.1. Model evaluation by metrics

The training and evaluation metrics for the five models are presented in Table 1. These metrics were computed using datasets aligned with each model's training configuration. From Model 8 (8 × 8 observation range) to Model 64 (64×64 range), a progressive improvement in accuracy is evident, demonstrating that larger observation ranges enhance model performance. For instance, Model 64 achieved a precision of 99.46%, significantly outperforming smaller-scale models. In contrast, the combined multi-scale model (8 + 16 + 32 + 64) exhibited the lowest precision (89.31%).

### 4.2. Model evaluation by visualization

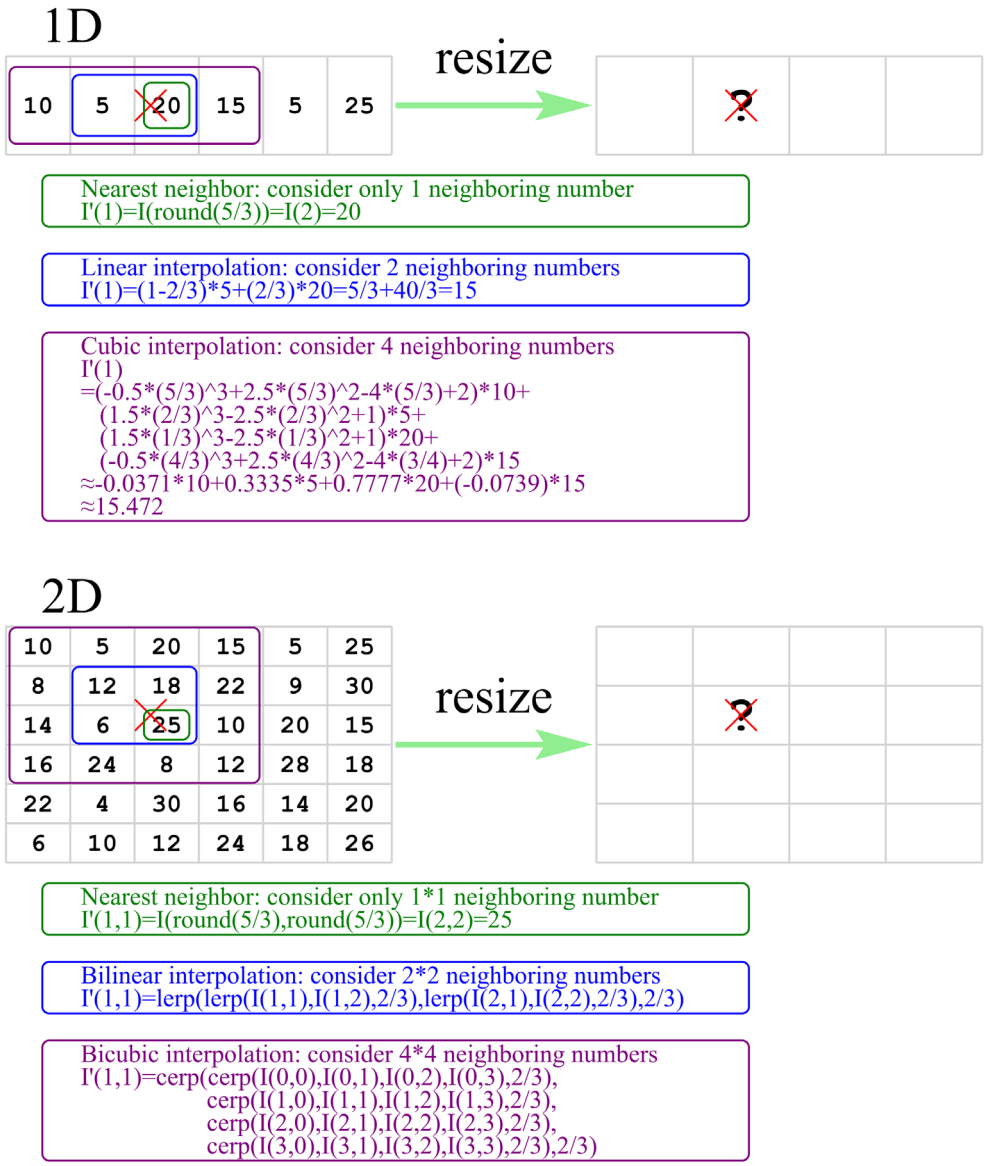



**Figure 5.** Demonstration of interpolation methods for one- and two-dimensional array resizing

**Table 1. Training and evaluation metrics of five models**

| Model | 8 | 16 | 32 | 64 | 8 + 16 + 32 + 64 |
|---|---|---|---|---|---|
| Context window | 8 × 8 | 16 × 16 | 32 × 32 | 64 × 64 | Combined |
| Training loss | 0.201 | 0.0936 | 0.0335 | 0.018 | 0.2415 |
| Training accuracy | 0.9198 | 0.9636 | 0.9877 | 0.9935 | 0.8977 |
| Validation loss | 0.3101 | 0.1308 | 0.0588 | 0.0359 | 0.2208 |
| Validation accuracy | 0.8774 | 0.9531 | 0.9844 | 0.9921 | 0.9086 |
| Precision | 0.9429 | 0.9811 | 0.9904 | 0.9946 | 0.8931 |
| Recall | 0.8436 | 0.9741 | 0.9963 | 0.9986 | 0.8015 |
| F1 Score | 0.8905 | 0.9776 | 0.9933 | 0.9966 | 0.8448 |

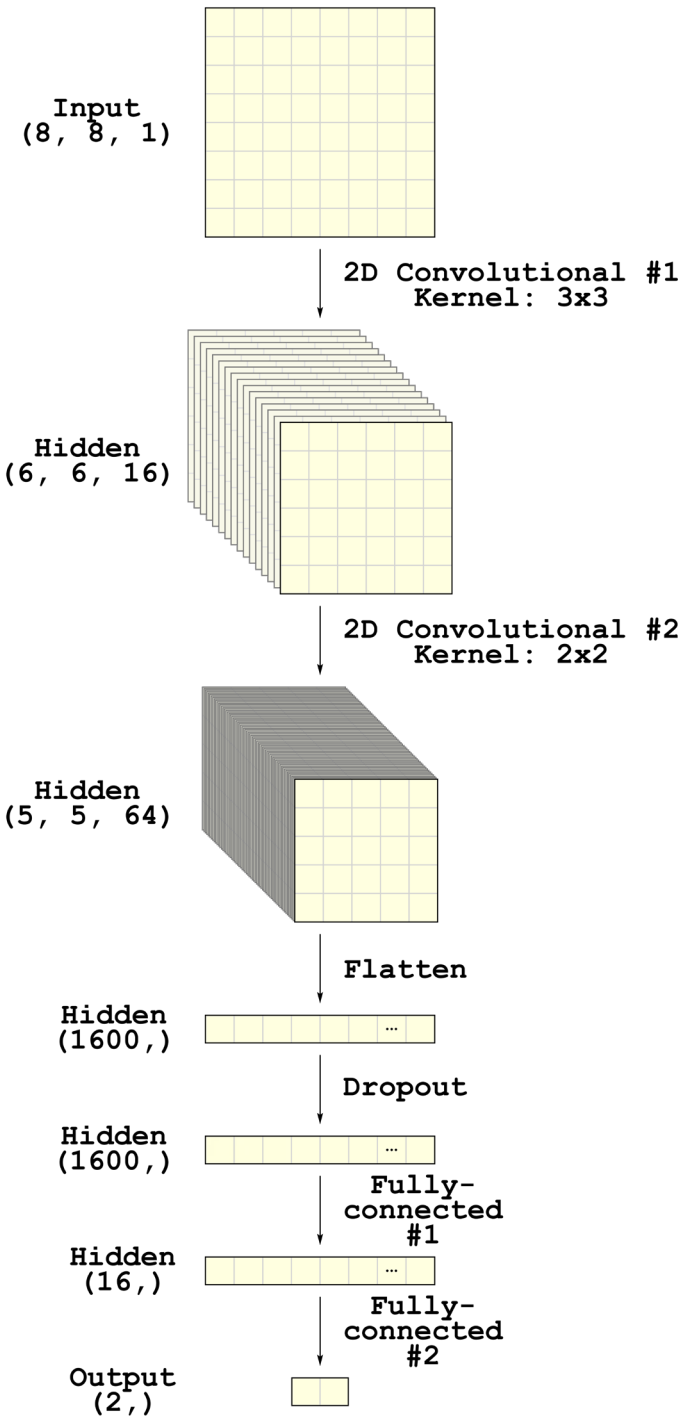



**Figure 6.** Lightweight deep convolutional architecture designed for this study. The model accepts an $8 \times 8 \times 1$ input matrix and outputs a binary classification result.

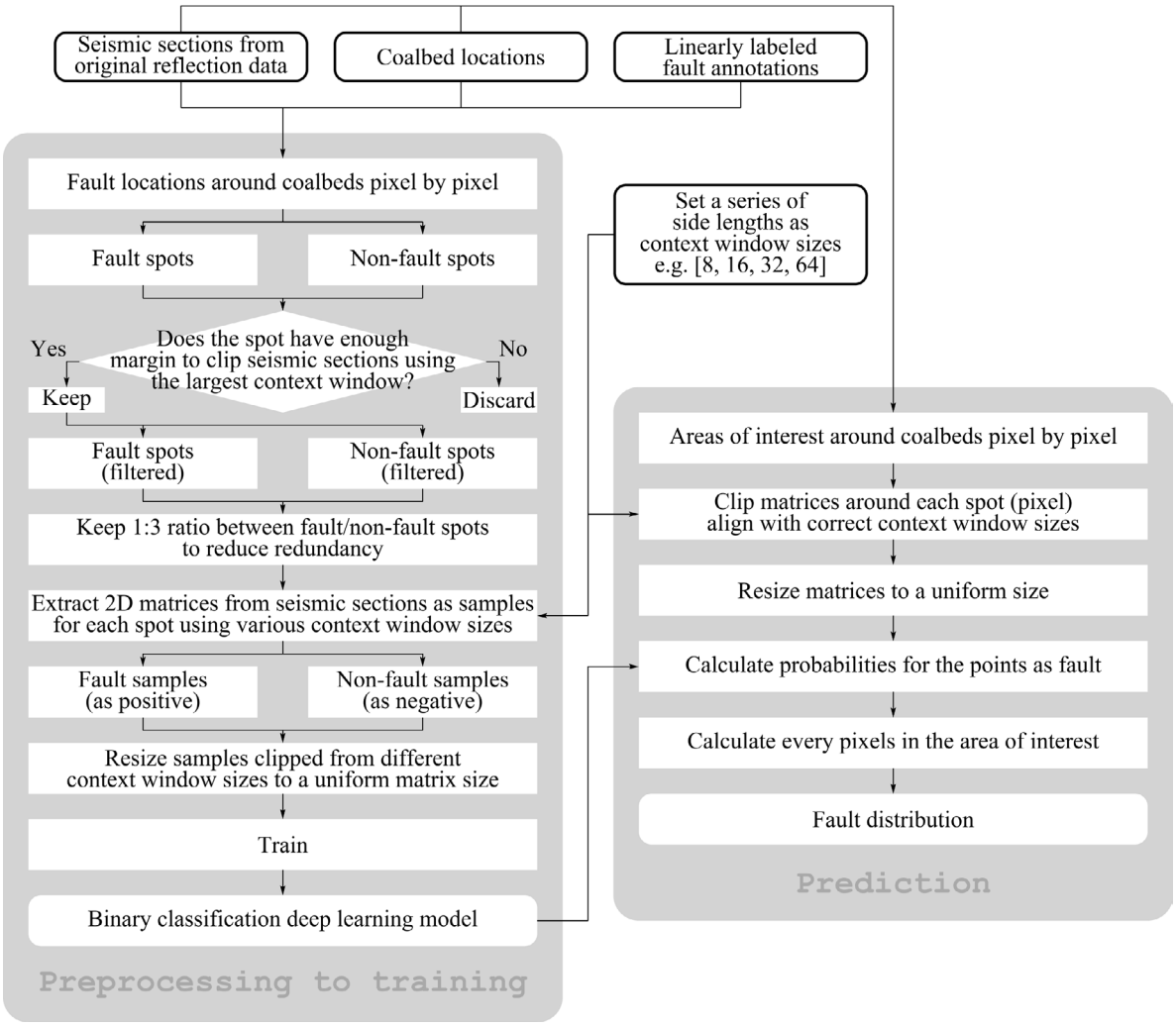



**Figure 7.** Proposed workflow, covering data preprocessing, model training, and prediction phases

Besides quantitative metrics, we also applied the trained models directly to seismic sections to compute fault/non-fault probabilities and generate corresponding distribution maps around the target coal seams. This visual assessment provides a practical evaluation of each model's performance. The resulting probability maps from applying different models to several test sections are shown in each subplot of Figure 8. Among them, the prediction result of the combined model (8+16+32+64) represents the averaged probability distribution computed from samples of all four different context window sizes fed simultaneously.

Visual assessment of the predicted fault distributions on seismic sections (Figure 8) indicates that Model 64 achieves the best continuity and closest agreement with the manual annotations. This observation aligns with the pattern of accuracy reflected in the quantitative metrics (Table 1).

### 4.3 Efficiency comparison of different input resolution

The central objective of this study is to improve computational efficiency in deep learning-based seismic fault interpretation. In principle, reducing the input resolution (*i.e.*, the size of the matrix fed into the model) can significantly decrease the number of network parameters, accelerate training, and reduce inference time during prediction. To quantitatively assess this efficiency gain, we compare the computational costs associated with different input resolutions. Specifically, we consider the scenario where models are designed natively for input sizes of $8 \times 8$, $16 \times 16$, $32 \times 32$, and $64 \times 64$ pixels, without the context-

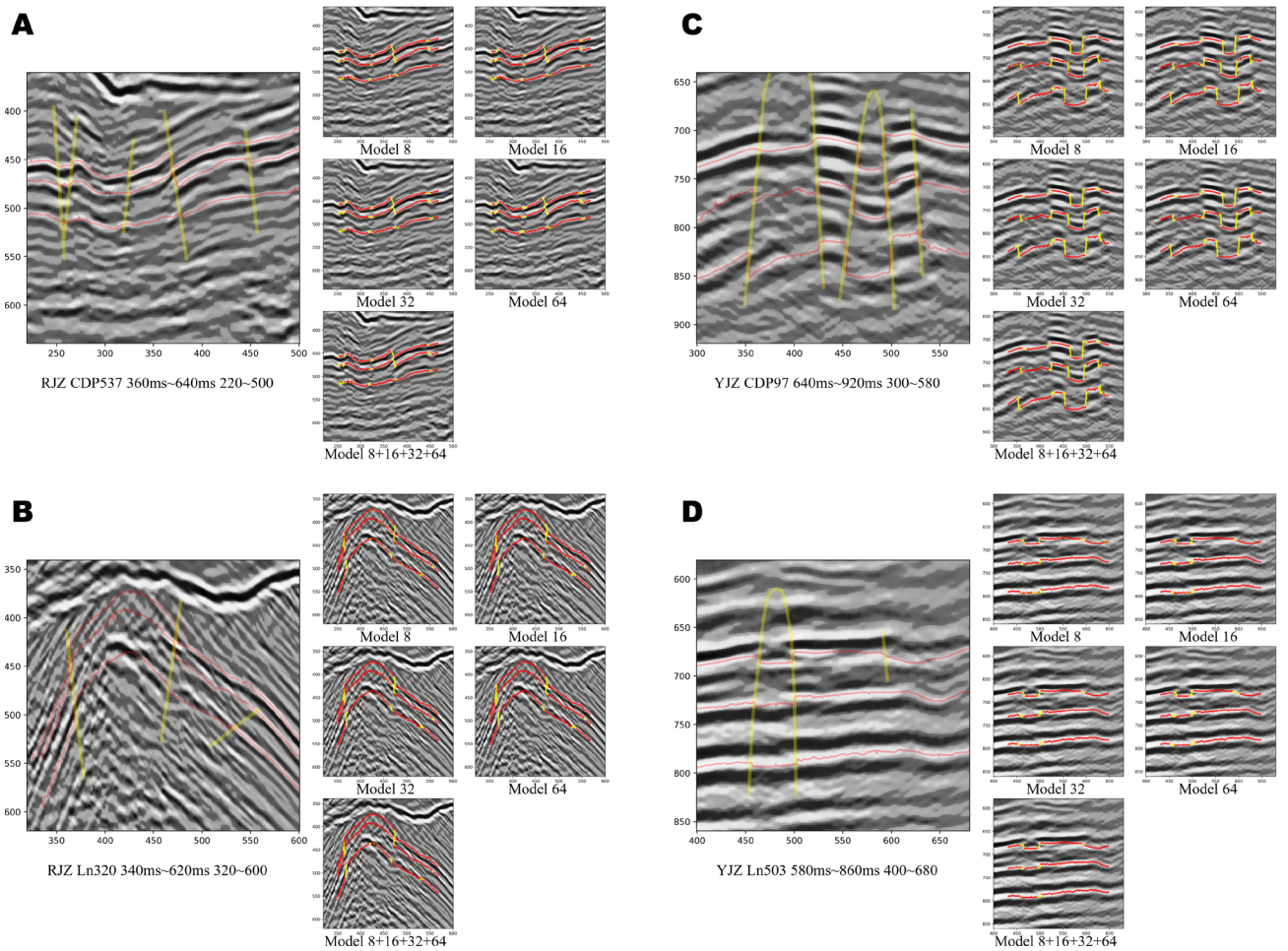



**Figure 8.** Four seismic sections from the two coal mine areas. (A) Renjiazhuang CDP537 360ms~640ms 220~500. (B) Renjiazhuang Ln320 340ms~620ms 320~600. (C) Yanjiazhuang CDP97 640ms~920ms 300~580. (D) Yanjiazhuang Ln503 580ms~860ms 400~680. Each panel includes original manual annotations (left) and fault distribution maps from the prediction results of various models (right).

window resizing strategy. The comparison focuses on three key metrics: the total number of trainable parameters in the CNN, the time required to complete 10,000 training epochs (the latter three did not complete the entire process, estimated based on the speeds of the previous epochs), and the average inference time per sample. The results are summarized in Table 2.

The 8 × 8 model requires only ~3k parameters, far fewer than the 64 × 64 model (~200k), leading to substantially faster training and inference. However, as shown in sections 4.1 and 4.2, using only a small native input window reduces accuracy. Our proposed strategy, extracting large context windows (*e.g.*, 64 × 64) and resizing them to a small fixed input (8 × 8), retains rich contextual information while operating within the highly efficient computational regime of the 8 × 8 model. This approach achieves high interpretation performance with significantly lower computational cost.

## 5. Discussion

The experimental results of this study demonstrate that the context window size is a critical factor influencing model performance. From 8 × 8 to 64 × 64, as the original spatial range used for sample extraction increases, the model's performance metrics show systematic improvement in both training and validation, with the largest window model ultimately achieving the highest accuracy and continuity in fault identification. This trend clearly indicates that even when all samples are uniformly resampled to the same low resolution (8 × 8) before being fed into the network, the scale information of the original geological structures they represent is still effectively preserved and conveyed to the model. A larger context window provides the model with richer spatial structure and contextual relationships,

**Table 2. Computational efficiency comparison for models with different native input resolutions**

| Input resolution | Model parameters | Training time (10,000 epochs) | Inference time per sample (ms) |
|---|---|---|---|
| 8 × 8 | ~3 k | ~15.6 h | ~0.05 |
| 16 × 16 | ~12.5 k | ~62 h | ~0.18 |
| 32 × 32 | ~50 k | ~240 h | ~0.65 |
| 64 × 64 | ~200 k | ~936 h | ~2.50 |

Note: Training times are estimated based on a single NVIDIA GTX 1650Ti GPU. Inference times are averaged over 100,000 samples.

enabling it to capture the overall morphology and seismic response characteristics of faults over a broader area, rather than isolated local anomalies. This confirms our core hypothesis: under the constraint of a fixed input resolution, adjusting the original observation window size for sample extraction allows for flexible control over the spatial context information of the input data, thereby significantly influencing the learning outcome and generalization capability of the deep learning model.

However, the combined model trained on samples from multiple window sizes (8 + 16 + 32 + 64) did not outperform the single best-window model, since its accuracy was notably lower. This suggests that simply mixing data from different scales did not automatically induce the anticipated "multi-scale feature fusion" advantage within the lightweight network architecture used here. A potential reason is that samples from different scales, after resizing, may possess inherently different feature distributions, and a simple classifier might struggle to learn a unified and robust cross-scale representation from them unbiasedly. More sophisticated network designs, such as incorporating feature pyramids or dedicated fusion modules, or employing staged, scale-specific training strategies, could be key to effectively leveraging multi-scale information.

Regarding computational efficiency, the strategy of "fixed small input resolution + variable large context window" proposed in this study presents a promising trade-off. All models were trained and predicted using a lightweight network designed for 8 × 8 input, which fundamentally constrains computational complexity. The results demonstrate that while maintaining a lightweight computational core, extracting information from larger original windows can yield performance comparable to or even surpassing that of traditional models with large inputs. This offers a new perspective for deploying deep learning models in practical resource-constrained environments.

The validation of this method on seismic data from two coal mines with distinct geological backgrounds indicates that the strategy based on pixelated samples and context window adjustment possesses a degree of geological generalizability. The models learned fault response patterns that are not confined to local features of a single area but are more universal. Compared to segmentation methods like U-Net, which require pixel-perfect annotations, this approach reduces the demand for annotation completeness, relying only on linear annotations. Compared to traditional machine learning methods based on handcrafted features, it avoids complex feature engineering. However, as an image classification approach, its output is discrete probability points. The continuity and morphological integrity of fault boundaries need to be reconstructed through post-processing of the probability maps, which remains a challenge when dealing with extremely complex or subtle fault systems.

Although deep learning models are often considered "black boxes", this study, through the controlled experiment of varying context windows, provides an indirect perspective for understanding model decision-making. The phenomenon of improving performance with larger windows suggests that model decisions heavily rely on regional context beyond local pixels. This insight guides future interpretability research to focus on analyzing the model's attention mechanisms toward features at different scales, thereby building a bridge from data scale to geological significance.

Therefore, the size of the context window is an important design parameter that connects the physical meaning of the data with the model's learning capacity. In the deployment of lightweight models that prioritize computational efficiency, the deliberate design and utilization of large context windows is an effective and practical strategy for enhancing model performance. Future work could focus on developing more intelligent architectures for multi-scale information fusion and further exploring the applicability

of this method to more complex geological structures and 3D seismic volume interpretation.

Furthermore, the lightweight modeling philosophy explored in this study, balancing computational efficiency with performance through context window adjustment and fixed low-resolution input, holds promising implications beyond seismic fault interpretation. Similar challenges of resource constraints and the need for efficient, scalable models are prevalent in other data-intensive geoscience domains, such as remote sensing image analysis and cloud detection. Recent advances have demonstrated the effectiveness of lightweight architectures in these areas. For instance, lightweight networks integrating multi-scale observation ranges and adaptive channel-spatial attention mechanisms have achieved high-precision cloud detection in remote sensing imagery, demonstrating feasibility in resource-constrained environments.[28] DecoupleNet introduced multi-branch feature decoupling and hybrid downsampling strategies to preserve small-scale objects in remote sensing imagery, achieving state-of-the-art performance with reduced computational overhead.[29] Similarly, Xiao *et al.*[30] proposed an enhanced inter-layer correlation framework for drone-based small object detection, demonstrating that adaptive resizing and feature separation significantly improve efficiency without compromising accuracy. The strategy proposed here, leveraging large contextual windows while standardizing inputs to a minimal dimension, could be adapted to other fields where data volume and model complexity are limiting factors. This cross-domain applicability highlights the broader relevance of our approach and encourages further exploration of efficient deep learning frameworks in geoscience and beyond.

In this study, we show that computational efficiency can be significantly improved with little sacrifice in training accuracy. However, this study also has some limitations:

(1) It lacks necessity to trade effectiveness for efficiency under the current dataset; however, for future applications with much larger datasets or more complex tasks, the proposed decoupling strategy offers a practical pathway to design more efficient models, where a controlled balance between efficiency and effectiveness may be meaningfully explored.
(2) The context window size was explored up to 64 × 64 pixels. While results show consistent improvement up to this size, whether further enlarging the window continues to enhance performance remains an open question. The choice of 64 × 64 was based on practical computational considerations and the scope of this work, which focused on validating the decoupling concept rather than pursuing maximum possible window sizes.
(3) The method was validated on seismic data from two coal mines with relatively clear fault structures. Its performance in highly complex geological settings (*e.g.*, salt domes, thrust belts, regions with structurally complex faults) has not been tested. Generalizability to such environments would require further investigation, possibly involving architectural adjustments or incorporation of domain-specific constraints.

## 6. Conclusion

This study systematically investigated the impact of context window size and input resolution on deep learning-based seismic fault interpretation, with a focus on computational efficiency. The core findings demonstrate that context window size is a crucial design parameter that effectively bridges the scale of geological data with model learning capacity, even when input dimensions are standardized. The key conclusions are:

(1) Under a fixed small input resolution (8 × 8 pixels), expanding the spatial context window (from 8 × 8 to 64 × 64) consistently and significantly improves model accuracy and the continuity of predicted faults. This confirms that richer contextual information, preserved despite downscaling, is critical for the model to learn robust, geologically-relevant features.
(2) Simply combining multi-scale samples into a single training set did not yield superior performance within our lightweight architecture, indicating that effective multi-scale fusion requires more sophisticated network designs or training strategies beyond simple data aggregation.
(3) The proposed strategy of "large context window + fixed small input resolution + lightweight network" establishes an efficient paradigm that maximizes computational performance per resource unit. By standardizing inputs to a minimal dimension (8 × 8), the network remains extremely compact, enabling significantly faster training and prediction cycles. Crucially, this architectural efficiency does not compromise accuracy, the model achieves superior interpretation quality by extracting richer geological context from larger original windows. This approach provides a practical and effective solution for deploying high-performance deep learning models in resource-constrained or time-sensitive scenarios.

## Acknowledgments

None.

## Funding

This work was supported by the National Key Research and Development Program of China (Grant number 2023YFB3211002) and the National Natural Science Foundation of China (Grant number 42274165).

## Conflict of interest

Suping Peng is one of the Editor-in-Chiefs of this journal, but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. The authors declare they have no competing interests.

## Author contributions

*Conceptualization*: Bowen Deng
*Data curation*: Guangui Zou, Suping Peng
*Methodology:* Bowen Deng
*Software*: Bowen Deng
*Supervision*: Suping Peng
*Writing–original draft*: Bowen Deng
*Writing–review & editing*: Guangui Zou, Suping Peng, Chengyang Han, Jingwen Xue

## Availability of data

Seismic data, codes and neural networks associated with this research are available from the corresponding author upon reasonable request.

## References

1. Karpatne A, Ebert-Uphoff I, Ravela S, Babaie HA, Kumar V. Machine Learning for the geosciences: Challenges and opportunities. *IEEE Trans Knowl Data Eng*. 2019;31(8):1544-1554.

   doi: 10.1109/TKDE.2018.2861006

2. Lary DJ, Alavi AH, Gandomi AH, Walker AL. Machine learning in geosciences and remote sensing. *Geosci Front*. 2016;7(1):3-10.

   doi: 10.1016/j.gsf.2015.07.003

3. Yu S, Ma J. Deep learning for geophysics: Current and future trends. *Rev Geophys*. 2021;59(3):e2021RG000742.

   doi: 10.1029/2021RG000742

4. Di H, Shafiq MA, AlRegib G. Seismic-fault detection based on multiattribute support vector machine analysis. In: *SEG Technical Program Expanded Abstracts 2017*. Proceedings of the 87th SEG Annual International Meeting; September 24-29, 2017; Houston, Texas, USA. Society of Exploration Geophysicists; 2017:2039-2044.

   doi: 10.1190/segam2017-17748277.1

5. Ren K, Zou G, Zhang S, Peng S, Gong F, Liu Y. Fault identification and reliability evaluation using an SVM model based on 3-D seismic data volume. *Geophys J Int*. 2023;234(1):755-768.

   doi: 10.1093/gji/ggad095

6. Han C, Zou G, Yeh HG, Gong F, Shi S, Chen H. Intelligent fault prediction with wavelet-SVM fusion in coal mine. *Comput Geosci*. 2025;194:105744.

   doi: 10.1016/j.cageo.2024.105744

7. Ren K, Zou G, Peng S, Yeh HG, Deng B, Ji Y. Fault identification based on the kernel principal component analysis-genetic particle swarm optimization-support vector machine algorithm for seismic attributes in the Sihe Coal Mine, Qinshui Basin, China. *Interpretation*. 2022;11(1):T59-T73.

   doi: 10.1190/int-2022-0039.1

8. Li W, Yang C, Sun D. Mining geophysical parameters through decision-tree analysis to determine correlation with tropical cyclone development. *Comput Geosci*. 2009;35(2):309-316.

   doi: 10.1016/j.cageo.2008.02.032

9. Li D, Peng S, Lu Y, Guo Y, Cui X. Seismic structure interpretation based on machine learning: A case study in coal mining. *Interpretation*. 2019;7(3):SE69-SE79.

   doi: 10.1190/INT-2018-0208.1

10. Ray A, Myer D. Bayesian geophysical inversion with trans-dimensional Gaussian process machine learning. *Geophys J Int*. 2019;217(3):1706-1726.

    doi: 10.1093/gji/ggz111

11. Brown WM, Gedeon TD, Groves DI, Barnes RG. Artificial neural networks: A new method for mineral prospectivity mapping. *Aust J Earth Sci*. 2000;47(4):757-770.

    doi: 10.1046/j.1440-0952.2000.00807.x

12. LeCun Y, Boser B, Denker JS, *et al*. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput*. 1989;1(4):541-551.

    doi: 10.1162/neco.1989.1.4.541

13. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60(6):84-90.

    doi: 10.1145/3065386

14. Bergen KJ, Johnson PA, de Hoop MV, Beroza GC. Machine learning for data-driven discovery in solid Earth geoscience. *Science*. 2019;363(6433):eaau0323.

    doi: 10.1126/science.aau0323

15. An Y, Du H, Ma S, *et al*. Current state and future directions for deep learning based automatic seismic fault interpretation: A systematic review. *Earth Sci Rev*. 2023;243:104509.

    doi: 10.1016/j.earscirev.2023.104509

16. Waldeland AU, Jensen AC, Gelius LJ, Solberg AHS. Convolutional neural networks for automated seismic

interpretation. *Lead Edge*. 2018;37(7):529-537.

doi: 10.1190/tle37070529.1

17. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; *Lecture Notes in Computer Science (9351)*. Proceedings of the MICCAI 2015, 18th International Conference; October 5-9, 2015; Munich, Germany. Springer; 2015:234-241.

    doi: 10.1007/978-3-319-24574-4_28

18. Huang J, Nowack RL. Machine learning using U-Net convolutional neural networks for the imaging of sparse seismic data. *Pure Appl Geophys*. 2020;177(6):2685-2700.

    doi: 10.1007/s00024-019-02412-z

19. Peters B, Haber E, Granek J. Neural networks for geophysicists and their application to seismic data interpretation. *Lead Edge*. 2019;38(7):534-540.

    doi: 10.1190/tle38070534.1

20. Wu X, Liang L, Shi Y, Fomel S. FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophysics*. 2019;84(3):IM35-IM45.

    doi: 10.1190/geo2018-0646.1

21. Cunha A, Pochet A, Lopes H, Gattass M. Seismic fault detection in real data using transfer learning from a convolutional neural network pre-trained with synthetic seismic data. *Comput Geosci*. 2020;135:104344.

    doi: 10.1016/j.cageo.2019.104344

22. Deng B, Zou G, Peng S, She J, Han C, Liu Y. An approach of 2D convolutional neural network–based seismic data fault interpretation with linear annotation and pixel thinking. *Geophys Prospect*. 2024;72(9):3350-3370.

    doi: 10.1111/1365-2478.13606

23. Li G, Li H, He S, Wang L. Multi-stage progressive network for seismic random noise suppression. *J Seism Explor*. 2025;34(1):43.

    doi: 10.36922/jse025240011

24. Wang W, Chen H, Chang D, Wang X, Wang S, Li D. Dual-branch dense network for seismic background noise elimination. *J Seism Explor*. 2025;34(5):53.

    doi: 10.36922/jse025290038

25. Parsania PS, Virparia PV. A comparative analysis of image interpolation algorithms. *Int J Adv Res Comput Commun Eng*. 2016;5(1):29-34.

    doi: 10.17148/IJARCCE.2016.5107

26. Jakhetiya V, Kumar A, Tiwari AK. A survey on image interpolation methods. In: Jusoff K, Xie Y, eds. *Volume 7546, Second International Conference On Digital Image Processing*. Proceedings of the Second International Conference on Digital Image Processing; February 26-28, 2010; Singapore. Society of Photo-Optical Instrumentation Engineers (SPIE); 2010:75461T.

    doi: 10.1117/12.855799

27. Han D. Comparison of commonly used image interpolation methods. In: Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013); Advances in Intelligent Systems Research Volume 34. 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013); March 22-23, 2013; Hangzhou, China. Atlantis Press; 2013:1556-1559.

    doi: 10.2991/iccsee.2013.391

28. Hu K, Zhang D, Xia M, Qian M, Chen B. LCDNet: Light-weighted cloud detection network for high-resolution remote sensing images. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2022;15:4809-4823.

    doi: 10.1109/JSTARS.2022.3181303

29. Lu W, Chen SB, Shu QL, Tang J, Luo B. DecoupleNet: A lightweight backbone network with efficient feature decoupling for remote sensing visual tasks. *IEEE Trans Geosci Remote Sens*. 2024;62:4414613.

    doi: 10.1109/TGRS.2024.3465496

30. Xiao Y, Xu T, Yu X, Fang Y, Li J. A lightweight fusion strategy with enhanced interlayer feature correlation for small object detection. *IEEE Trans Geosci Remote Sens*. 2024;62:1-11.

    doi: 10.1109/TGRS.2024.3457155

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# GeoSSL: A geology-aware self-supervised framework for fault detection in 3D seismic data

**Mengge Wang[1]*, Xinrong Hu[2], Junyu Zhang[3], and Haofa Lin[4]**

[1]School of Mathematics and Computer Science, Hanjiang Normal University, Shiyan, Hubei, China
[2]School of Computer and Artificial Intelligence, Wuhan Textile University, Wuhan, Hubei, China
[3]School of Computer Science & Technology, Beijing Institute of Technology, Beijing, China
[4]School of Artificial Intelligence, Wuhan Technology and Business University, Wuhan, Hubei, China

## Abstract

Fault identification is a critical task in 3D seismic interpretation, directly influencing the efficiency and accuracy of reservoir characterization and drilling decisions. However, traditional methods rely heavily on manual experience and high-quality annotated data, making it difficult to adapt to the demands of processing massive amounts of seismic data. To address this, we introduce a novel self-supervised learning (SSL) paradigm specifically designed for geological feature learning, which for the first time unifies masked voxel reconstruction, 3D patch contrastive learning, and multimodal attribute joint contrast into a coherent multi-task pre-training framework. This framework is uniquely tailored to leverage the spatial continuity and physical attributes of seismic data, enabling the model to learn transferable, geologically meaningful representations without any manual labels. It leverages unlabeled seismic data to learn geologically meaningful feature representations and leverages a transfer learning mechanism to achieve high-precision fault identification with small sample sizes. Experiments on multiple public and field-measured datasets, including F3 and SEAM Phase I, demonstrate that this method achieves key metrics such as intersection over union (IoU) and F1-scores of 0.76 and 0.87, respectively, significantly outperforming traditional attribute analysis (IoU = 0.49) and supervised deep learning models (IoU = 0.72). Furthermore, the method remains robust in areas with low signal-to-noise ratios (average confidence > 85%) and consistently estimates fault strike and dip (average absolute error ≤ 3.5°). This research provides an effective solution for reducing reliance on manual annotation and improving the reliability of automated fault interpretation in complex tectonic areas.

***Keywords:*** Self-supervised learning; 3D seismic interpretation; Fault identification; Contrastive learning; Multimodal pre-training

***Corresponding author:**
Mengge Wang
(wangmengge@hjnu.edu.cn)

**Citation:** Wang M, Hu X, Zhang J, Lin H. GeoSSL: A geology-aware self-supervised framework for fault detection in 3D seismic data. *J Seismic Explor*. 2026;35(2):025510127. doi: 10.36922/JSE025510127

**Received:** December 20, 2025

**Revised:** February 24, 2026

**Accepted:** March 10, 2026

**Published online:** April 22, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Commonly used in oil and gas exploration and geological hazard prediction, 3D seismic data provide rich subsurface structural information, making them highly valuable for resource assessment and risk management.[1] Fault identification and interpretation are crucial components of seismic data interpretation. Accurate characterization of fault systems not only impacts the effectiveness of reservoir evaluation but also directly impacts

the success rate of drilling deployment and the profitability of exploration and development.[2] With advances in seismic acquisition technology and the rapid growth of data volumes, traditional manual interpretation methods are no longer able to meet the demands for efficient and detailed interpretation.[3] There is an urgent need to develop automated and intelligent fault identification technologies.

The present mainstream fault identification methods can be classified into three categories: manual interpretation, traditional attribute analysis, and supervised deep learning. Manual interpretation relies on the empirical judgment of geologists. While highly accurate, it is inefficient and highly subjective, often resulting in significant inter-interpreter variability. Traditional seismic attribute analysis methods—such as coherence volumes and curvature attributes—enhance fault responses through mathematical transformations, improving interpretation efficiency to a certain extent.[4] However, these methods often rely on manually designed features, have limited generalization capabilities under complex geological conditions, and are ineffective in identifying weak faults. Supervised deep learning methods can automatically learn fault features and perform well on large amounts of labeled data. However, their performance is highly dependent on high-quality training samples. The high cost and time required to label earthquake faults severely restrict the widespread application of these methods in practical applications.

Self-supervised learning (SSL), an emerging paradigm in machine learning, can automatically learn effective representations from unlabeled data, offering a new approach to addressing the scarcity of labeled data.[5] In geophysics, SSL holds enormous potential. 3D seismic data, with its rich spatial structure and physical laws, provide an ideal data foundation for SSL.[6-8] However, the application of this technology in seismic interpretation is still in its infancy. In fault identification, many research gaps remain, including how to design effective self-supervised pre-training tasks, how to transfer learned features to downstream tasks, and how to ensure the reliability of models in practical applications.

The key innovations and contributions of this work are fourfold, directly addressing the aforementioned gaps:

(1) A novel, multi-task SSL framework is proposed that synergistically combines masked reconstruction (for local texture), 3D contrastive learning (for global context), and multimodal attribute joint comparison (for physical consistency). While individual SSL tasks have been explored in geophysics, to the best of our knowledge, this work represents one of the first attempts to synergistically combine them within a coherent framework specifically tailored for 3D seismic fault interpretation. The originality of our approach lies particularly in the multimodal joint contrast mechanism, which, unlike generic contrastive learning applied to a single data type, explicitly leverages the complementary geophysical information from multiple seismic attributes (such as amplitude, phase, coherence). By contrasting these different views of the same subsurface location, the model is forced to learn representations that are invariant to attribute-specific noise and more faithful to the underlying geology, thereby enhancing feature robustness and transferability.

(2) A boundary-sensitive feature enhancement module is introduced driven by seismic gradient information. This module, a significant departure from generic SSL architectures that treat all voxels equally, explicitly guides the model to attend to fault discontinuities by weighting features based on the local gradient magnitude. This design is not merely an architectural addition but a geology-aware inductive bias that prioritizes learning from regions most indicative of faults, thereby improving the delineation of subtle and complex fault boundaries.

(3) An efficient transfer learning pipeline is developed, that demonstrates the superior versatility of our pre-trained features across multiple high-value interpretation tasks, such as fault segmentation, classification, and strike/dip estimation, and achieves high accuracy with very few labeled samples.

(4) Extensive, rigorous validation is provided not only on standard metrics but also on geological consistency, uncertainty quantification, and practical application (such as fault seal analysis), establishing a new benchmark for reliability in automated seismic interpretation.

This study focuses on the identification of structural faults (such as normal, reverse, strike-slip) within post-stack 3D seismic amplitude data and derived attributes. The proposed framework is designed for regions with moderate to good data quality and is validated in both siliciclastic (F3) and complex salt-influenced (SEAM) settings. Its applicability to areas with extremely low signal-to-noise ratio (SNR), heavily karstified strata, or volcaniclastic sequences requires further investigation and potential adaptation.

This study is organized as follows: Section 2 reviews related research; Section 3 details the dataset, preprocessing pipeline, and experimental platform; Section 4 elaborates on the design and implementation of the self-supervised model; Section 5 discusses the fault identification and

interpretation method; Section 6 presents experiments and results analysis; Section 7 verifies the model's generalization and practical application; Section 8 discusses its innovations and limitations; and Section 9 concludes the paper and outlines future research directions. Through systematic theoretical exploration and experimental validation, this study aims to advance 3D seismic fault identification technology toward a more intelligent and efficient direction.

To ensure conceptual clarity and delineate the scope of application, we explicitly define the methodological boundaries and key terminology. The proposed framework is developed for structural interpretation of conventional post-stack 3D seismic amplitude data and derivative attributes, targeting seismically resolvable fault discontinuities. It is architecturally optimized for structural fault identification rather than lithological fault discrimination, sub-resolution fracture characterization, or direct fluid prediction. Critical computational constructs are operationally defined: the fault probability volume denotes a 3D confidence field where each voxel value represents the model's estimated likelihood of fault presence, while pseudo-labeling refers to the iterative process of generating training labels from high-confidence model predictions on unlabeled data.

## 2. Review of related research

The application of SSL techniques in 3D seismic data interpretation and fault identification has attracted considerable attention in recent research. These methods aim to improve the accuracy and generalization of fault detection without over-reliance on labeled datasets. Such as a recent study highlighted the role of self-supervised pre-training in earthquake fault identification, demonstrating improvements in 3D segmentation.[9] This approach exploits the inherent structure of seismic data to learn meaningful representations, which facilitate more accurate fault delineation.

The development of specialized models such as SeisCoDE further demonstrates progress in this area. SeisCoDE is a foundational model for 3D seismic interpretation that employs a novel SSL framework that combines seismic signal processing and attribute analysis to preserve structural integrity.[10] These models highlight the importance of SSL in capturing seismic features that are crucial for fault interpretation. Similarly, contrastive reconstruction-based methods utilize self-supervised representation learning to improve the generalization of fault detection networks, enabling more robust identification across diverse seismic datasets.[11] Attention-enhanced architectures have also contributed to this field, with models like Attention FaultFormer combining 3D Convolutional Neural Network (CNN) and Transformer modules in a self-supervised pre-training paradigm to improve fault identification performance.[12] These models benefit from the Transformer's ability to capture long-range dependencies and contextual information, which is crucial for accurate fault interpretation in complex seismic volumes.

In addition, semi-supervised methods that combine SSL with lightweight modules have been proposed to improve existing Segment Anything Models for earthquake fault detection, further demonstrating the versatility of SSL techniques.[13] End-to-end neural networks like ResACEUnet embody the trend of models that can directly interpret 3D seismic data without human intervention, leveraging SSL to enhance feature extraction and fault delineation.[14]

The consensus in the literature suggests that machine learning, especially SSL, is significantly advancing seismic interpretation because it enables models to effectively learn from unlabeled data. This shift not only reduces the reliance on large, labeled datasets but also improves the robustness and accuracy of fault detection systems.[15-18] Collectively, these studies highlight the great potential of SSL-based methods in advancing the field of 3D seismic fault identification and interpretation.

Despite these promising advances, critical gaps remain in applying SSL to seismic fault identification. First, existing methods often rely on a single pre-training task (such as reconstruction or contrastive learning), failing to holistically capture both the detailed local texture (essential for fault edges) and the global semantic context (essential for fault continuity) of 3D seismic volumes. Second, the rich, physically meaningful information contained in multiple seismic attributes is underexploited in a joint, contrastive manner. Third, there is a lack of a streamlined pipeline that effectively transfers SSL-learned features to diverse downstream interpretation tasks (segmentation, parameter estimation) with minimal supervision. Finally, the robustness and geological plausibility of SSL-based predictions in low-SNR areas and complex structures require systematic validation.

## 3. Research data and experimental platform

### 3.1. Description of 3D seismic dataset

This study used three different types of 3D seismic datasets to train and validate the fault identification algorithm, covering both simulated data and real exploration data. The F3 Demo dataset comes from the North Sea region

of the Netherlands and contains typical inclined bed and delta sedimentary structures, providing rich seismic phase and reflection surface annotation information. SEAM Phase I is an authoritative numerical simulation dataset for deepwater saline areas with precise physical parameters and synthetic seismic data, which is suitable for performance testing of the algorithm in complex geological environments. The actual data in the exploration area comes from real oil and gas exploration projects, including manually annotated fault lines and wellbore verification records, which provide an important basis for evaluating the effectiveness of the algorithm in practical applications. The dataset comparison is shown in Table 1.

These three datasets have their own characteristics in spatial resolution, temporal sampling rate and data scale, and together constitute a comprehensive and representative test benchmark.

### 3.2. Data preprocessing process

Both the raw 3D seismic data and annotations require preprocessing to adapt to SSL and downstream fault identification tasks. The process is as follows:

The volume data is normalized and augmented. The raw amplitude volume is $D(i,j,k)$, where $i$ and $j$ are the inline and crossline indices, respectively, and $k$ is the time sampling point index.[19] Normalization is defined as:

$$\tilde{D}(i,j,k)=\frac{D(i,j,k)-\mu}{\sigma} \tag{1}$$

where $\mu=\mathbb{E}[D]$ and $\sigma=\sqrt{\mathbb{E}[(D-\mu)^2]}$, and the statistics are calculated over the entire data volume or on training sub-volume patches.

This study employed multiple data augmentation techniques to improve the model's generalization and robustness. Random flipping reverses the data in either the inline or crossline directions, increasing the diversity of the training data. Random rotation transforms slices by multiples of 90 degrees, effectively simulating geological structures in different orientations. Random cropping samples patches of varying sizes (64 × 64 × 64) from the original volume data, ensuring that the model can handle geological features at varying scales. Gaussian noise

**Table 1. Comparison of 3D seismic fault identification datasets**

| Dataset name | Region/Tectonic type | Spatial resolution (inline × crossline spacing) | Time sampling rate/Time window | Attribute and annotation information | Number of samples/ voxels |
|---|---|---|---|---|---|
| F3 Demo | Netherlands North Sea / including inclined bed, deltaic | 25 m × 25 m | 1.848 ms sampling step/0–1200 ms | Contains multiple seismic phases, several reflection surfaces (horizons), wellbore data | ~386.9 km² area; voxel size approximately 400 × 400 × 600 |
| SEAM Phase I | Simulation models of deepwater saline areas and other complex structures | 20 m × 20 m × 10 m (after depth/ time conversion) | 8 ms/full-depth model | Density, P-S wave velocity, synthetic lithology facies, and synthetic seismic data | Area approximately 35 km × 40 km × 15 km depth |
| Actual data of the exploration area | Oil and gas exploration area | 25 m × 25 m | 2 ms/0–2000 ms | Fault lines have been marked, and some wellbore penetration records have been recorded. | Voxel size: 400 × 400 × 1000, approximately $1.6 \times 10^8$ voxels |

addition, implemented after data normalization, enhances the model's robustness to data noise by introducing normally distributed noise $n(i,j,k) \sim \mathcal{N}(0,\sigma_n^2)$, with a noise intensity $\sigma_n$ set to 0.01. These augmentation strategies work together to significantly expand the scope of the training dataset and improve the model's adaptability and stability under diverse geological conditions.

To rigorously evaluate model performance, we established a clear split for the three datasets. For the F3 Demo dataset, 80% of the patches were randomly selected for training, 10% for validation to tune hyperparameters, and the remaining 10% for testing. The SEAM Phase I dataset was held out entirely for cross-dataset generalization testing and did not participate in any training phase. The real-field exploration dataset was divided based on spatial location: a complete subsurface structural block (approximately 70% of the volume) was used for training, while an adjacent, independent structural block (approximately 30% of the volume) was used for testing. This spatial split ensures independence between the training and test sets, preventing data leakage. All reported performance metrics are calculated based on the respective test sets.

To address the prevalent class imbalance in fault detection tasks (where fault voxels are significantly outnumbered by non-fault voxels), we adopted a combined strategy during training. First, we incorporated Dice Loss into our loss function (as shown in **Equation 11**), which is inherently robust to class imbalance. Second, during data loading, we implemented an online hard example mining technique that dynamically increases the sampling weight of patches containing fault voxels, ensuring the model adequately learns from these scarce but critical features.

In addition to class imbalance, we also implemented systematic noise processing and quality control to address data quality issues in the original F3 data. The noise primarily stems from shallow seafloor spurious waves, out-of-band high-frequency noise, and amplitude drift. We employed dip-steered median filtering, with a filter radius set to the neighborhood of two inline/crossline lines surrounding the trace, effectively eliminating random noise while preserving fault edge features. A bandpass filter with a passband range of 5–60 Hz, adjusted based on the spectral analysis results of each dataset, effectively removed low-frequency seismic background trends and high-frequency noise interference. Edge trimming specifically addresses undersampled traces at volume boundaries or in areas of low coverage, ensuring uniform data quality. Wellbore alignment correction accurately converted seismic fault travel times into depth information through time-depth calibration of seismic-well data, precisely aligning reflection surfaces with corresponding markers in the wellbore, significantly improving the spatial accuracy of the data.

For data annotation quality control, we established a rigorous quality assurance system. For actual fault line or surface annotations, experts conducted geometric consistency reviews to ensure accuracy. For horizon or facies labels, the system checked for gaps or overlaps between labels to ensure annotation integrity. Consistency verification was performed between pixel-level and voxel-level annotations to ensure consistency across different scales. In terms of signal enhancement, we employ a variety of techniques to improve the discernibility of weak faults, including the generation of auxiliary channels using fault-enhancing attributes such as amplitude differences, variance attributes, and similarity attributes. These auxiliary channels, as an important component of the model input, significantly enhance the model's ability to detect weak fault features. The final training input is fed into the network in the form of blocks, typically of size P × P × T, such as 64 × 64 × 64 or 32 × 32 × 128, where T represents the number of sampling points in the time dimension.

### 3.3. Experimental platform and hardware environment

This study constructed a high-performance computing platform to ensure the reproducibility and computational efficiency of large-scale 3D seismic data fault identification experiments. This hardware platform utilizes a heterogeneous architecture design combining graphics processing units (GPUs) and central processing units (CPUs), specifically optimized for the memory-intensive and massively parallel computing requirements of seismic volume data processing. Four NVIDIA RTX 4090 GPUs provide powerful parallel computing capabilities suitable for self-supervised pre-training and large-scale volume training tasks. A dual-socket Intel Xeon Silver 4214 CPU system provides 48 logical processing cores for data pre-processing and auxiliary computing. 512GB of large-capacity memory ensures simultaneous loading and processing of multiple large seismic volume data blocks, while the high-speed NVMe SSD storage system significantly improves data read efficiency and reduces I/O bottlenecks. A 100 Gbps high-speed network interconnect supports rapid data transfer between multiple nodes, providing the necessary network bandwidth for distributed training. This hardware configuration provides a solid infrastructure for the efficient execution of deep learning algorithms on 3D seismic data.

This study employed advanced deep learning

frameworks and tool environments to ensure experimental reliability and reproducibility. The deep learning frameworks primarily used were PyTorch (Version ≥ 1.11) and TensorFlow (Version ≥ 2.6). Cross-validation was used to ensure algorithm correctness and consistent results. Code management and experimental logging were performed using the Git version control system, combined with visualization tools such as Weights & Biases or TensorBoard, enabling complete tracking of experimental progress and visual monitoring of results. Data formatting supported the industry-standard SEG-Y format (Via the segyio library) as well as the efficient HDF5 format, ensuring accurate reading and efficient storage of seismic data. Visualization employed a multi-layered toolkit: Matplotlib and Seaborn for 2D slices and statistical charts, and Mayavi and Paraview for interactive visualization of 3D volumetric data, providing comprehensive visual support for result analysis and interpretation.

The training process employed carefully optimized parameter configuration and run strategies. The batch size setting carefully considered the balance between computing resources and model performance. On a single GPU, batches of 2 were used to process $64 \times 64 \times 64$ blocks, and multi-GPU parallel training enabled a total batch size of 8. The optimizer uses the Adam or AdamW algorithm, with an initial learning rate set to $lr0 = 1 \times 10^{-4}$. Learning rate decay strategies such as cosine decay or step decay are employed, dynamically adjusting the learning rate based on the number of training epochs or steps to ensure stable convergence during training. The training cycle is tailored to the specific task. Self-supervised pre-training typically requires 200–300 epochs to fully learn data features, while downstream tasks such as fault segmentation and classification typically achieve optimal performance with 100–150 epochs. To ensure reproducibility, a fixed random seed (such as seed = 42) was used across all experiments to eliminate the influence of random factors and ensure the reliability and comparability of the research conclusions.

By combining public and real-world datasets, a systematic preprocessing strategy, and a powerful computing platform, this research provides a solid foundation for subsequent self-supervised pre-training and fault identification tasks.

## 4. Model design based on SSL

To fully exploit the fault geometry and amplitude characteristics in 3D seismic data, a multi-task framework based on SSL is constructed. This framework first pre-trains large-scale unlabeled data through tasks such as mask reconstruction, 3D patches contrast learning, and joint comparison of multimodal attributes. The pre-trained features are then transferred to the downstream task of fault identification and interpretation. The overall concept

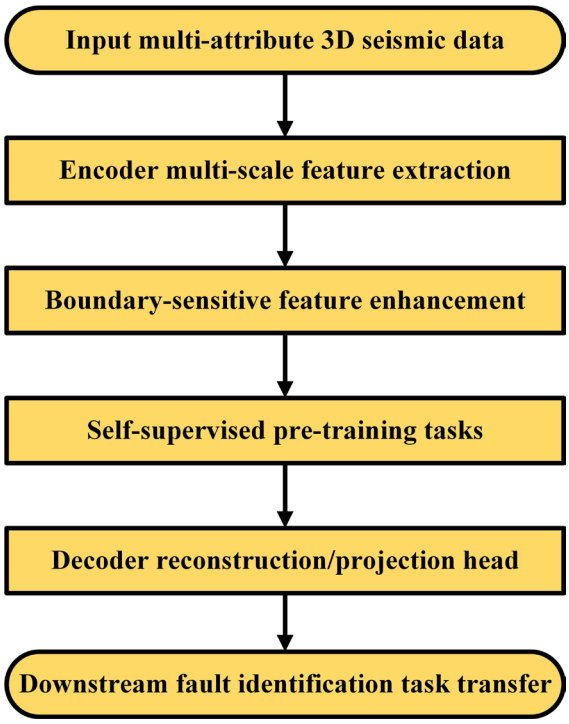



**Figure 1.** Model architecture flow chart

is shown in Figure 1.

To provide a clear technical blueprint of the proposed framework, we present a detailed schematic of the model architecture and data processing flow in Figure 2. This figure illustrates the two-phase pipeline: (1) the multi-task self-supervised pre-training on unlabeled seismic

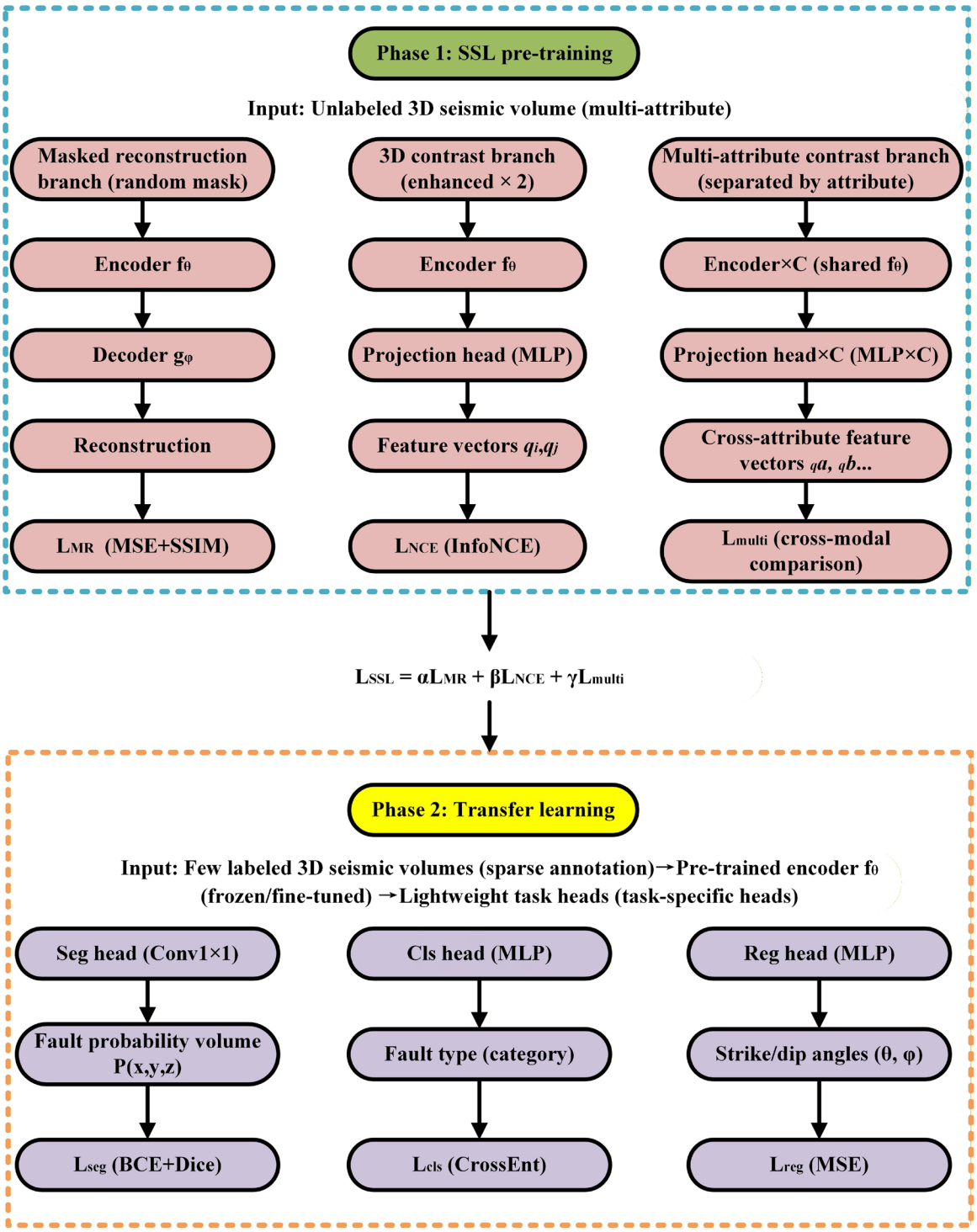



**Figure 2.** Detailed model architecture and data flow diagram

volumes, and (2) the transfer learning for downstream fault interpretation tasks.

Figure 2 detailed architecture and data flow of the proposed SSL-based fault interpretation framework. (1) Self-supervised pre-training phase: An unlabeled 3D seismic volume (or multi-attribute volume) is processed. It undergoes data augmentation and is fed into three parallel SSL tasks: Masked Reconstruction (with a masking module and encoder-decoder), 3D Patch Contrastive Learning (with two augmented views processed by a shared encoder and projection head), and Multimodal Joint Contrast (with different attribute channels processed and compared). Their losses are combined to update the encoder. (2) Transfer learning phase: The pre-trained encoder is frozen or fine-tuned, and connected to lightweight, task-specific heads (such as segmentation, classification, regression heads) for supervised training on a small set of labeled data, yielding fault probability volumes, classifications, or geometric attributes.

In terms of architectural design, a 3D convolutional encoder-decoder integrated with a multi-scale feature extraction module is employed to capture sensitive information about fault boundaries.[20] Let the original normalized input patch be $\mathrm{X} \in \mathbb{R}^{C\times D\times H\times W}$ , where $C$ is the number of channels ( $C>1$ for multi-attribute input), and $D,H,W$ are the depth, row, and column dimensions, respectively. The encoder $f_\theta$ maps this to latent features $Z$ , which are then reconstructed by the decoder $g_\phi$ or other self-supervised outputs $\hat{\mathrm{X}}$ .

$$\mathrm{Z}=f_\theta(\mathrm{X}),\hat{\mathrm{X}}=g_\phi(\mathrm{Z}) \tag{2}$$

The encoder consists of multiple 3D convolutional blocks (kernel size 3 × 3 × 3) with residual connections.[21] The multi-scale module introduces different receptive fields $r$ :

$$\mathrm{Z}_r=\mathrm{Conv3D}_r(\mathrm{X}),\mathrm{Z}=\sum\nolimits_{r\in\{1,2,4\}}\mathrm{Z}_r \tag{3}$$

where $r$ represents the dilation rate or pooling scale. Features of different scales are concatenated in the channel dimension and fed into the fault boundary sensitivity module. This module uses gradient magnitude or similarity as weights to enhance fault neighborhood features:

$$\mathrm{W}(i,j,k)=\exp\left(-\frac{\|\nabla \mathrm{X}(i,j,k)\|^2}{2\sigma^2}\right),Z_{\text{fault}}=\mathrm{W}\odot\mathrm{Z} \tag{4}$$

This allows the fault edge to be highlighted in the latent space while suppressing the background continuous reflection.

Table 2 lists the main layers of the model architecture and parameter configuration.

This network utilizes an encoder-decoder architecture and is optimized for feature extraction and fault probability prediction from 3D volume data. The encoder layer progressively extracts feature representations at different levels through multi-scale convolution and downsampling operations. Block 2 employs downsampling to halve the spatial dimension while increasing the number of channels. The boundary-sensitive module specifically enhances the ability to extract fault edge features through a gradient weighting mechanism. The decoder layer progressively restores spatial resolution through upsampling and convolution operations, and the final output layer uses a sigmoid activation function to generate a fault probability map. Techniques such as residual connections and batch normalization are incorporated into the network design to ensure training stability and effective feature transfer.

During the self-supervised pre-training phase, three tasks are designed:

(1) Masked reconstruction:

Randomly mask a portion of the voxels in the input block (mask ratio $m$, such as 30%–50%) and only input the remaining portion $\mathrm{X}_{\text{visible}}$ into the encoder. The decoder needs to reconstruct the complete block $\hat{\mathrm{X}}$ . The loss function is the reconstruction means squared error (MSE):[22-24]

$$\mathcal{L}_{\text{MR}}=\frac{1}{|\Omega_m|}\sum\nolimits_{(i,j,k)\in\Omega_m}\left(\hat{\mathrm{X}}(i,j,k)-\mathrm{X}(i,j,k)\right)^2 \tag{5}$$

where $\Omega_m$ is the set of masked voxels. We also introduce the structural similarity (SSIM) term to improve the reconstruction quality:[25]

$$\mathcal{L}_{\text{SSIM}}=1-\mathrm{SSIM}\left(\hat{\mathrm{X}},\mathrm{X}\right) \tag{6}$$

The total loss $\mathcal{L}_{\text{MR+SSIM}}=\mathcal{L}_{\text{MR}}+\lambda\mathcal{L}_{\text{SSIM}}$ .

(2) 3D patch-wise contrastive learning:

The two views $\mathrm{Z}_i,\mathrm{Z}_j$ generated at different positions or in different enhancement methods are fed into the projection head $h(\cdot)$ to obtain vectors $\mathrm{q}_i,\mathrm{q}_j$ . The InfoNCE loss is used:[26]

$$\mathcal{L}_{\text{contrast}}=-\log\frac{\exp\left(\cos\left(\mathrm{q}_i,\mathrm{q}_j\right)/\tau\right)}{\sum_{k=1}^{N}\exp\left(\cos\left(\mathrm{q}_i,\mathrm{q}_k\right)/\tau\right)} \tag{7}$$

where $\cos\left(\mathrm{q}_i,\mathrm{q}_j\right)=\frac{\mathrm{q}_i^{\top}\mathrm{q}_j}{\|\mathrm{q}_i\|\|\mathrm{q}_j\|}$, and $\tau$ is the temperature

**Table 2. Model architecture and parameter configuration**

| Module/Layer name | Input size | Convolution kernel/ stride | Number of channels | Activation function | Other notes |
|---|---|---|---|---|---|
| Encoder Block1 | 64 × 64 × 64 × C | 3 × 3 × 3 /1 | 32 | ReLU | 3D convolution + batch normalization |
| Encoder Block2 (Down) | 32 × 32 × 32 × 32 | 3 × 3 × 3 /2 | 64 | ReLU | Multi-scale convolution/Dilated convolution |
| Encoder Block3 | 16 × 16 × 16 × 64 | 3 × 3 × 3 /1 | 128 | ReLU | Residual connection |
| Boundary Sensitive Mod. | 16 × 16 × 16 × 128 | - | 128 | Sigmoid | Gradient weighted enhancement of fault edges |
| Decoder Block1 (Up) | 32 × 32 × 32 × 64 | 3 × 3 × 3/1 | 64 | ReLU | Upsampling + convolution |
| Decoder Block2 (Up) | 64 × 64 × 64 × 32 | 3 × 3 × 3/1 | 32 | ReLU | Upsampling + convolution |
| Output Layer | 64 × 64 × 64 × 1 | 1 × 1 × 1/1 | 1 | Sigmoid | Reconstruction/Prediction of fault probability |

Abbreviation: ReLU: Rectified linear unit.

coefficient. This contrastive learning paradigm, which encourages the model to learn invariant representations by maximizing agreement between different views of the same data, has proven effective not only in seismic interpretation but also in other geoscience-related fields, such as remote sensing, where consistency reasoning has been explored to enhance object detection performance.[27]

(3) Multi-attribute pretraining:

The selection of seismic attributes for multimodal input is grounded in their complementary geophysical roles in fault characterization. We deliberately choose four core attributes: amplitude, instantaneous phase, instantaneous frequency, and similarity (coherence). This combination provides a balanced and physically meaningful representation for SSL.

Amplitude provides the primary signal of subsurface impedance contrast. Instantaneous phase delivers robust sensitivity to lateral discontinuities, crucial for identifying fault-related reflector breaks even in low-amplitude zones. Instantaneous frequency supplies lithological and stratigraphic context, as its abrupt variations are often associated with fault zones due to wave diffraction effects. Similarity (coherence) is a mathematically defined

attribute engineered to highlight discontinuities directly. Its inclusion creates a powerful bootstrapping effect, guiding the model to implicitly learn fault-sensitive features from the raw data. Collectively, this attribute suite balances direct structural signals with nuanced contextual cues. Their joint contrastive learning forces the model to discover the underlying geological structures that are consistently expressed across these diverse representations, thereby yielding more generalizable and robust feature embeddings.

The input contains $C \geq 4$ attribute channels, such as amplitude, instantaneous phase, instantaneous frequency, and similarity, and the representations of different attributes are compared in the latent space. The loss function can be defined as multimodal cross-channel comparison:[28]

$$\mathcal{L}_{\text{multi}} = \sum_{a \neq b} -\log \frac{\exp\left(\sin\left(\mathbf{q}_i^{(a)}, \mathbf{q}_j^{(b)}\right)/\tau\right)}{\sum_k \exp\left(\sin\left(\mathbf{q}_i^{(a)}, \mathbf{q}_k^{(b)}\right)/\tau\right)} \tag{8}$$

The total self-supervision loss is the weighted sum of three:

$$\mathcal{L}_{\text{SSL}} = \alpha \mathcal{L}_{\text{MR+SSIM}} + \beta \mathcal{L}_{\text{contrast}} + \gamma \mathcal{L}_{\text{multi}} \tag{9}$$

The hyperparameters $\alpha$, $\beta$, and $\gamma$ control the contribution of each self-supervised task to the overall pre-training objective. Through empirical experimentation on the validation set, these weights were set to $\alpha = 1.0$, $\beta = 0.5$, and $\gamma = 0.8$. This configuration balances the three tasks, giving primary importance to masked reconstruction while maintaining significant contributions from contrastive learning and multimodal joint contrast. A sensitivity analysis was performed by varying each parameter within the range $[0.2, 2.0]$ while holding others fixed; the results showed that the model performance (measured by downstream fault segmentation IoU) remained stable within $\pm 10\%$ of the chosen values, demonstrating robustness to moderate hyperparameter variations.

The schematic diagram shown in Figure 3 clearly shows the key links of mask reconstruction and 3D patches comparison learning in the self-supervised pre-training process.

Figure 3A shows a randomly masked 3D seismic volume. A certain percentage of voxels in the original amplitude data are randomly masked. This forces the network to understand local structural and textural features with only partial information. Figure 3B shows the model's reconstruction of the masked region after

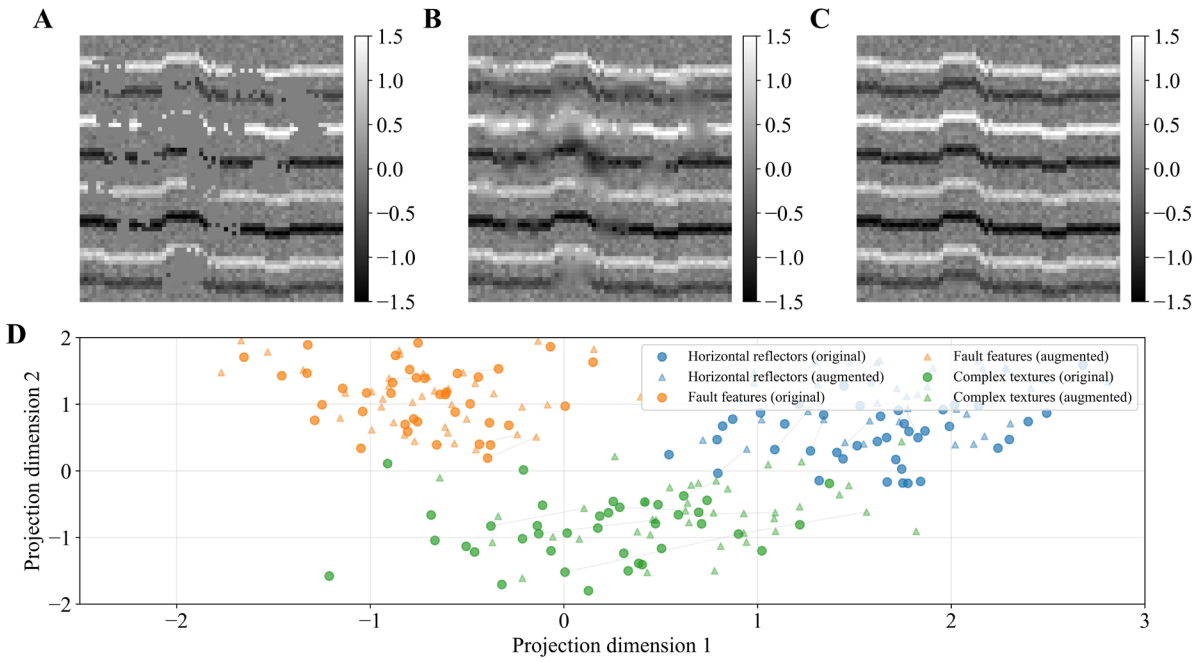



**Figure 3.** Self-supervised learning framework for seismic data: masked reconstruction and contrastive feature learning. (A) Input seismic section with patch-based masking preserving geological continuity, (B) reconstructed seismic data showing recovered reflector layers and fault features, (C) original complete seismic section with horizontal reflectors and structural fault for comparison, and (D) contrastive learning projection space where different geological feature types form distinct clusters, demonstrating effective feature representation learning.

training. The fault edges, amplitude transitions, and small-scale seismic phase features of the reconstructed volume are well restored, highly consistent with the original data in terms of geometry. Figure 3C shows the complete input volume, used to compare the consistency between the model reconstruction and the real data. Figure 3D illustrates the clustering of the projected space of the same seismic volume after different enhancements. The blue and red dots represent the two enhanced views, respectively, which naturally form clusters in the projected space. This demonstrates that the contrastive learning branch can map different views of the same semantic meaning to adjacent regions, thereby improving the discriminability and robustness of the features. This figure vividly demonstrates that after our multi-task joint optimization during the pre-training phase, the model not only fills in missing information but also clusters structural semantics in the latent space, facilitating subsequent fault identification.

The encoder features $Z$ obtained through the above pre-training can be transferred to the downstream fault identification task, including fault classification, segmentation, and strike/dip prediction.[29] We introduce a lightweight segmentation head for the fault segmentation task:

$$\hat{\mathrm{Y}} = \sigma\left(\mathrm{Conv3D}_{1\times1\times1}\left(\mathrm{Z}\right)\right) \tag{10}$$

where $\hat{\mathrm{Y}}$ is the fault probability volume. The loss function is binary cross entropy + Dice loss:

$$\mathcal{L}_{\text{fault}} = -\sum_{i,j,k}\left[y_{ijk}\log\hat{y}_{ijk} + \left(1-y_{ijk}\right)\log\left(1-y_{ijk}\right)\right] + 1 - \frac{2\sum y\hat{y}}{\sum y + \sum \hat{y}} \tag{11}$$

For fault strike/dip prediction, we regress the fault strike $\theta$ and dip $\phi$ on $Z$ to minimize the mean square error:[30,31]

$$\mathcal{L}_{\theta,\phi} = \frac{1}{N}\sum_{n=1}^{N}\left[(\hat{\theta}_n - \theta_n)^2 + (\phi_n - \phi_n)^2\right] \tag{12}$$

Crucially, downstream tasks operate in two distinct modes, determined by label availability.

*Supervised Fine-tuning with Limited Labels:* For surveys with sparse manual interpretations (such as 5%–10% of the volume, as used in the main experiments described in Section 6), we directly fine-tune the pre-trained encoder using a lightweight task head and standard supervised losses (**Equations 11** and **12**). This setup leverages pre-training to achieve high performance with minimal annotation.

*Unsupervised Extraction via Pseudo-Labeling:* For completely unlabeled data, we employ a teacher-student framework with consistency regularization. The teacher model generates an initial fault probability volume. High-confidence predictions (threshold $\tau > 0.8$) are converted into pseudo-labels after morphological cleaning. The student model is trained on strongly augmented versions of the input to match these pseudo-labels via a consistency loss (such as MSE). The teacher's weights are updated via an exponential moving average of the student's weights. The total loss combines this consistency loss with the self-supervised loss (**Equation 9**) to prevent confirmation bias. For small samples or unlabeled areas, we use pseudo-labeling + consistency regularization to achieve automatic fault extraction, clustering self-supervised features or predicting fault probabilities on unlabeled data through a teacher-student network.

Table 3 systematically demonstrates the self-supervised pre-training strategies and their transfer performance in the downstream fault identification task.

The study employed three different self-supervised pre-training methods: mask reconstruction pre-training (MSE = 0.012), 3D patches contrastive learning pre-training (InfoNCE = 0.45), and multimodal attribute joint contrastive pre-training (InfoNCE = 0.38). These methods all successfully learned meaningful feature representations to varying degrees. The pre-training phase fully utilized large amounts of unlabeled data (60,000–100,000 patches), while the downstream fine-tuning phase required only a small amount of labeled data (2,000–5,000 samples), significantly reducing the annotation cost. Multimodal attribute joint contrastive pre-training achieved the best InfoNCE loss (0.38) using four attribute channels, demonstrating that the fusion of multi-attribute information facilitates the learning of richer feature representations. The pre-trained features were successfully transferred to a variety of downstream tasks, including fault segmentation (IoU = 0.83), fault classification (Acc = 0.92), and geological parameter regression (MAE = 3.5°/2.1°), demonstrating the strong generalization capability of the pre-trained model. Particularly noteworthy is that even without annotations, pseudo-labeling combined with consistency training still achieves an IoU of 0.77, demonstrating the practicality of this approach in real-world applications.

Figure 4 shows the effect of transferring pre-trained features to the downstream segmentation task

Figure 4A shows the original seismic section, which shows continuous reflectors, relatively weak fractures, and background noise. Figure 4B shows the manually interpreted fault annotations, with the fault locations highlighted. The outlines were drawn point by point by geological interpretation experts based on seismic

**Table 3. Self-supervised pre-training and downstream task transfer configuration and performance**

| Task types | Number of training samples (patch) | Mask ratio m | Number of comparison views | Number of attribute channels | Migrating downstream tasks | Evaluation metrics |
|---|---|---|---|---|---|---|
| Mask reconstruction pre-training | 100,000 | 0.4 | - | 1 | Encoder pre-training | MSE = 0.012 |
| 3D patches contrastive learning pre-training | 80,000 | - | 2 | 1 | Encoder pre-training | InfoNCE = 0.45 |
| Multimodal attribute joint contrastive pre-training | 60,000 | - | 4 | 4 | Encoder pre-training | InfoNCE = 0.38 |
| Fault segmentation transfer (with annotations) | 5,000 | - | - | ¼ | Segmentation head | IoU = 0.83 |
| Fault classification transfer (with annotations) | 3,000 | - | - | ¼ | Classification head | Acc = 0.92 |
| Fault strike/Dip prediction (with annotations) | 2,000 | - | - | ¼ | Regression head | MAE = 3.5°/2.1° |
| Small sample automatic extraction (without annotations) | - | - | - | ¼ | Pseudo-labeling + consistency | IoU = 0.77 |

Abbreviations: Acc: Accuracy; InfoNCE: Information noise-contrastive estimation; IoU: Intersection over union; MAE: Mean absolute error; MSE: Mean squared error.

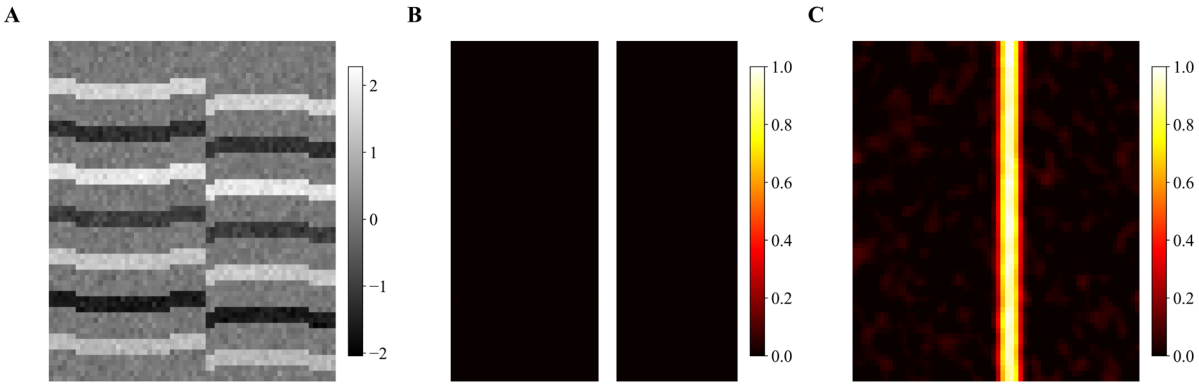



**Figure 4.** Transfer learning for fault segmentation using pre-trained seismic features. (A) Input seismic section displaying horizontal reflector sequences and fault-displaced layers; (B) Expert-annotated fault labels showing precise fault plane delineation; (C) Model-predicted fault probability map using pre-trained self-supervised features, demonstrating improved fault detection accuracy compared to traditional methods.

reflection discontinuities, which are highly subjective. Figure 4C shows the fault probability volume obtained after transferring the self-supervised pre-trained model. The color intensity indicates the model's confidence in the existence of the fault. Compared to manual annotations, the model predicts more consistent fault locations, effectively captures fractures in weak reflection areas, and displays a continuous geometric pattern of strike and dip. This demonstrates that the deep features extracted by pre-training are highly sensitive to fault boundaries. Even in the absence of sufficient labeled samples, high-precision fault probability volumes can be generated, providing more stable and automated support for seismic interpretation.

Through this design, our model acquires rich structural and amplitude characterization capabilities on large-scale unlabeled 3D seismic data, and then realizes fault classification, segmentation and geometric attribute prediction through a lightweight downstream head network. It can automatically extract fault structure even in small samples or unlabeled areas, providing a high-precision, automated tool for seismic interpretation.

## 5. Fault identification and interpretation methods

For fault extraction and segmentation, we propose a 3D fault imaging pipeline based on self-supervised features. The model first performs feature mapping on the seismic volume to obtain the fault-sensitive feature vector $f_i \in \mathbb{R}^d$ for each voxel. The model then outputs the fault probability volume $P(x,y,z)$ through a 3D convolutional decoder, which is expressed as:

$$P(x,y,z) = \sigma\left(W_2 * \phi\left(W_1 * F(x,y,z)\right)\right) \tag{13}$$

where $F(x,y,z)$ is the deep feature tensor output by the self-supervised model, $W_1, W_2$ are trainable convolution kernels, $\phi$ is a nonlinear activation function, and $\sigma$ is a sigmoid function that constrains the output to [0,1] to represent the fault probability.

Figure 5 shows a comparison of 3D tomographic images generated based on self-supervised features.

The original seismic section in Figure 5A primarily displays a continuous reflection layer system, where fault locations are often obscured by noise, impedance variations, and multi-stage sedimentary overlays, resulting in blurred boundaries that are difficult to accurately identify. The fault probability volume slice in Figure 5B utilizes a depth model to estimate the probability of fault presence for each voxel. The results show that the fault boundary is highlighted with a high probability value, the continuity of the background layer system is effectively suppressed, and the spatial location and geometric outline of the fault are more clearly defined. The 3D maximum projection in Figure 5C further integrates the high-confidence fault region of the entire volume data into a 2D image, making the fault network structure clear and facilitating macroscale interpretation. A comparison of the three models demonstrates that the SSL model can extract stable fault features in complex backgrounds, achieving continuous visualization from local slices to the entire 3D space.

Automatic estimation of multi-scale fault strike and dip utilizes principal direction analysis of the feature tensor.[32]

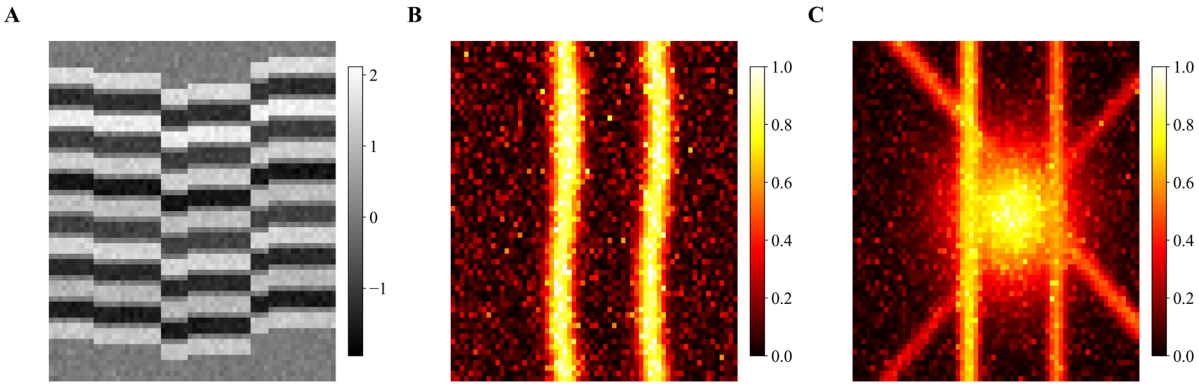



**Figure 5.** 3D fault imaging and visualization: from seismic data to fault network characterization. (A) Original seismic section showing continuous horizontal reflectors with fault-displaced layers and structural discontinuities, (B) fault probability slice displaying computed likelihood of fault presence with enhanced fault corridor delineation, and (C) 3D maximum projection visualization revealing complex fault network architecture including vertical, diagonal, and curved fault systems in three-dimensional space.

Structural tensor decomposition is performed on the local region of the probability volume:

$$T = \begin{bmatrix} \langle g_x^2 \rangle & \langle g_x g_y \rangle & \langle g_x g_z \rangle \\ \langle g_x g_y \rangle & \langle g_y^2 \rangle & \langle g_y g_z \rangle \\ \langle g_x g_z \rangle & \langle g_y g_z \rangle & \langle g_z^2 \rangle \end{bmatrix} \tag{14}$$

where $g_x, g_y, g_z$ are the gradients of the probability volume in three directions. The notation ‹·› denotes local spatial averaging within a small analysis window centered at each voxel, typically implemented using a Gaussian or uniform kernel. By eigendecomposition, the main eigenvector $v_{max}$ is obtained, and the fault strike $\theta$ and dip angle $\phi$ are obtained:

$$\theta = \arctan 2\left(v_y, v_x\right), \phi = \arctan\left(\frac{v_z}{\sqrt{v_x^2 + v_y^2}}\right) \tag{15}$$

Understanding the geometric properties of fault networks, such as their branching patterns and hierarchical organization, connects to fundamental research in graph theory and tree structures. For instance, exact combinatorial methods have been developed to analyze subtree properties in tree-like networks,[33] providing theoretical foundations that could inform future topological analysis of fault systems extracted from seismic data.

Figure 6 shows the automatic estimation results of multi-scale fault strike and dip.

Figure 6A superimposes local directional arrows on the fault probability volume. These arrows, derived from gradient field and tensor analysis, indicate the direction of local fault extension and reflect the geometric orientation of the fault in space. Figure 6B shows the thermal distribution of strike, with different color ranges representing different azimuth angles. The figure clearly shows how the fault's horizontal extension direction varies with position, forming a regular strike zone. Figure 6C presents the corresponding dip distribution. The color variation reveals the vertical tilt of the fault. Some areas with larger dips may represent recently active faults, while more gradual areas may represent relics of earlier tectonic activity. Overall, the multi-scale estimation not only provides a quantitative description of the fault's spatial morphology but also lays the foundation for subsequent tectonic evolution analysis and mechanical modeling.

To quantify the model's ability to interpret fault geometry and activity, we selected eight representative faults in different tectonic regions and extracted their geometric parameters, including length, average dip, displacement, and activity period. Table 4 systematically presents the geometric characteristics and activity parameter calculations for the eight faults, fully demonstrating the excellent performance of the self-supervised feature-driven algorithm in weakly reflective regions.

Fault lengths range from 6.8 to 16.7 km, with F3 being the largest (16.7 km) and F5 the smallest (6.8 km), demonstrating the diversity of faults in the study area. Dip parameters range from 48° to 63°, typical of normal or high-angle reverse faults, with F2 having the highest dip (63°) and F5 the lowest (48°). Displacement varies significantly, with the maximum reaching 150 m (F3) and the minimum only 15 m (F5), reflecting the varying activity intensities of the faults. Activity period estimates indicate that F5 has the longest activity period (4.2 Ma), while F3 has the shortest (1.8 Ma), likely reflecting the regional tectonic evolution. Confidence levels are generally

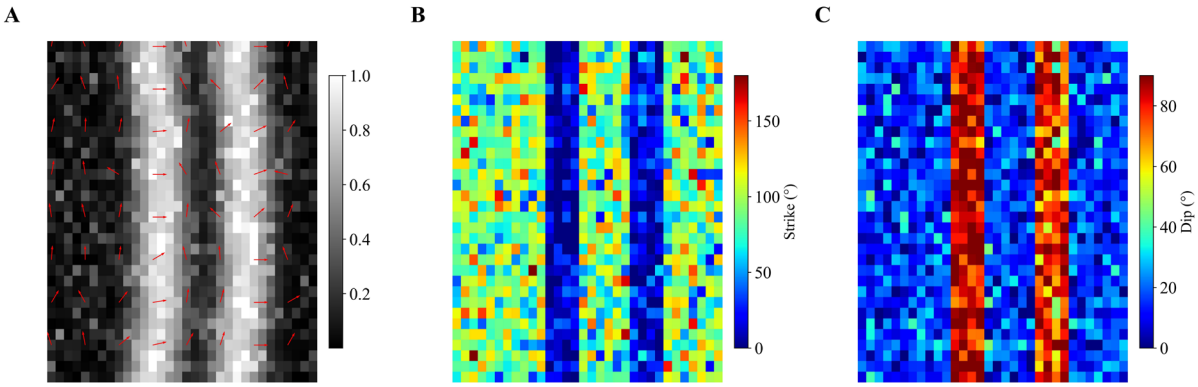



**Figure 6.** Multi-scale automatic estimation of fault strike and dip attributes from seismic data. (A) Fault probability overlaid with local orientation vectors showing directional analysis based on fault geometry and structural trends; (B) Strike distribution heatmap displaying azimuthal variations with fault-related near-north-south trends and bedding-related east-west orientations; (C) Dip angle distribution showing steep fault planes (80–90°) contrasted with gentle sedimentary layer dips (5–25°), demonstrating structural geological differentiation.

**Table 4. Comparison of different fault geometry and activity parameters**

| Fault number | Length (km) | Average inclination (°) | Maximum displacement (m) | Minimum displacement (m) | Activity period (Ma) | Confidence (%) | Toward consistency (%) |
|---|---|---|---|---|---|---|---|
| F1 | 12.3 | 58 | 120 | 30 | 2.1 | 92 | 88 |
| F2 | 8.5 | 63 | 85 | 20 | 3.5 | 89 | 86 |
| F3 | 16.7 | 55 | 150 | 40 | 1.8 | 95 | 91 |
| F4 | 10.4 | 60 | 90 | 25 | 2.9 | 90 | 87 |
| F5 | 6.8 | 48 | 70 | 15 | 4.2 | 85 | 80 |
| F6 | 14.1 | 62 | 110 | 35 | 2.4 | 93 | 89 |
| F7 | 9.9 | 57 | 95 | 20 | 3.0 | 88 | 84 |
| F8 | 11.6 | 59 | 105 | 28 | 2.6 | 91 | 86 |

high (85%–95%), demonstrating the algorithm's ability to reliably detect faults even in areas of weak reflection, with the highest confidence level (95%) for F3. The strike consistency index ranges from 80% to 91%, demonstrating the algorithm's superiority in extracting fault continuity, particularly in areas with complex geological conditions, where it maintains high spatial consistency.

Figure 7 demonstrates the integration of fault interpretation with sedimentological context.

In Figure 7, different colors are used to distinguish sedimentary facies types (sandstone, shale, limestone, and mixed facies) with an overlay of fault probability volumes (hot colors), creating a coupled visualization of faults and sedimentary environments. Some faults penetrate multiple sedimentary facies units, exhibiting distinct cross-layer cutting characteristics (such as cutting through sandstone, shale, and limestone), which suggests strong regional tectonic control. Other faults, however, are primarily confined to a single sedimentary facies, potentially related

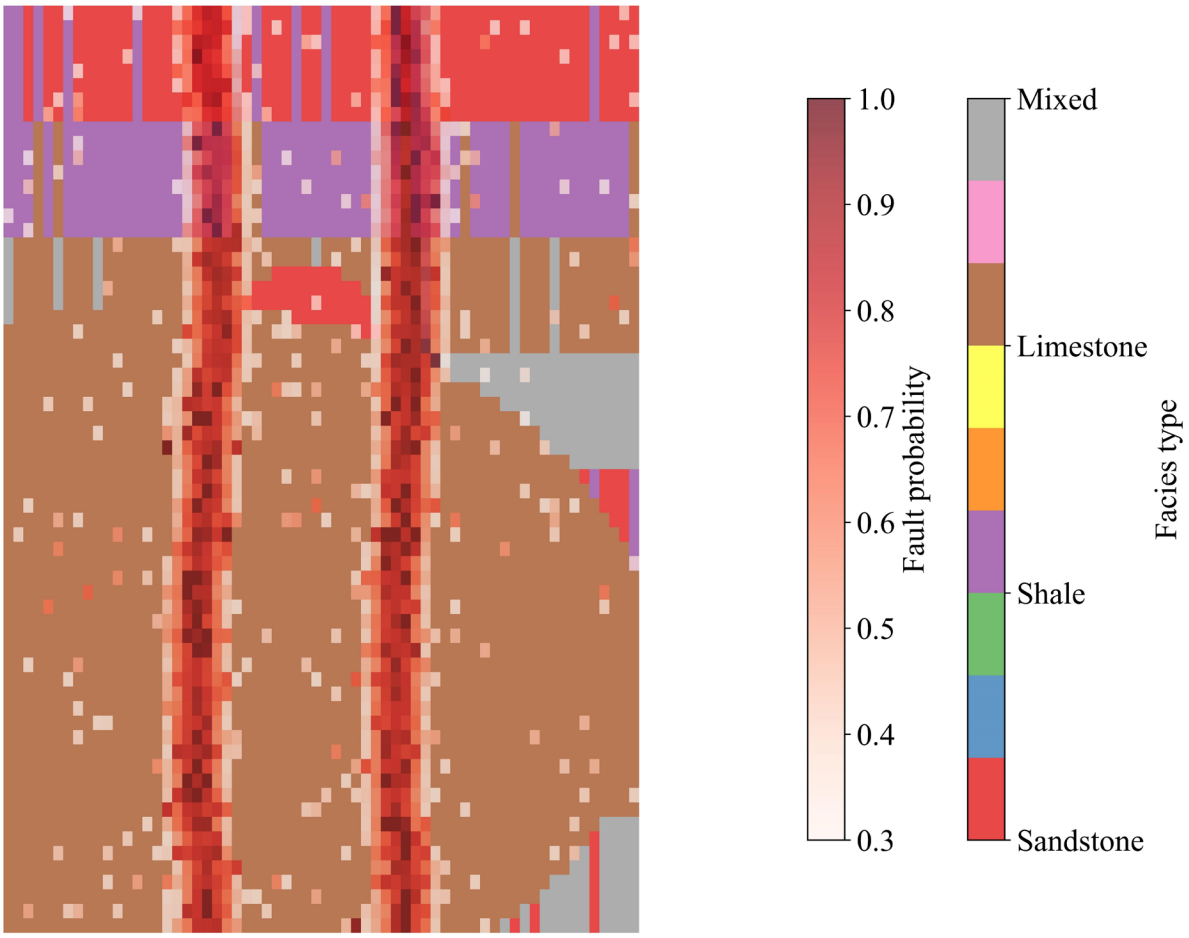



**Figure 7.** Integration of fault probability volumes with sedimentary facies analysis for reservoir characterization. The two color bars on the right define the two datasets: The right color bar corresponds to the four sedimentary facies types distinguished using a categorical color scheme, while the left color bar indicates the fault probability, ranging from low (transparent) to high (dark red).

to local stress field adjustments during sedimentation. This overlay interpretation helps understand the impact of faults on sedimentary facies distribution, reservoir distribution, and fluid migration pathways.

To enhance the reliability of the results, this study also conducted uncertainty quantification and credibility analysis of the fault identification results. Monte Carlo Dropout was used to calculate the variance of the fault probability volume during the inference phase using multiple forward propagations, defining the uncertainty metric:

$$U(x,y,z) = \mathrm{Var}\left(P_1(x,y,z), P_2(x,y,z), \ldots, P_N(x,y,z)\right) \quad (16)$$

where $P_N(x,y,z)$ is the fault probability value predicted by the $N$-th sampling. A larger $U$ indicates a lower confidence in the model at that voxel location.

In practical application, the workflow can be adaptive: For new seismic surveys with no prior interpretation, the model first operates in the unsupervised pseudo-labeling mode (as described in Section 4) to generate an initial fault probability volume and geometric attribute estimates. Interpreters can then selectively validate and correct high-uncertainty regions or key structural areas (as indicated by Table 5), creating a small set of high-confidence labels. These labels are then used to switch to the supervised fine-tuning mode, rapidly refining the model specifically for the local geology of the new survey. This hybrid human-in-the-loop workflow maximizes efficiency while minimizing manual annotation effort. Table 5 summarizes the statistical results for different seismic regions. Within tectonic zones with high SNRs, the model predicts higher confidence and lower standard deviations for faults. In contrast, in areas with complex backgrounds, ambiguous fault features, or

high noise, the uncertainty of the model output increases significantly, providing a quantitative reference for manual interpretation.

As shown in Table 5, the mean probability values for areas A1 and A2 are both greater than 0.8, with low standard deviations (<0.12), indicating that the model's predictions in these areas are highly consistent and reliable. In contrast, in areas C and D, particularly the unlabeled region of D2, the mean probability is only approximately 0.37, with a high standard deviation of 0.28 and an uncertainty index close to 0.8, indicating that the model lacks sufficient discriminative power in these areas. The results must be verified through manual interpretation and additional geological data. This approach, based on a combined analysis of mean probability and uncertainty, provides quantifiable metrics for subsequent exploration deployment and risk control, facilitating more robust fault interpretation in complex geological settings.

## 6. Experiment and results analysis

In this study, to systematically evaluate a 3D seismic fault identification method based on SSL, we designed multiple comparative experiments and introduced multi-dimensional performance evaluation metrics for analysis. These comparative experiments included traditional fault detection methods based on seismic attributes (such as

**Table 5. Statistics of mean and standard deviation of uncertainty in fault interpretation in different regions**

| Area number | Geological background description | Number of sample points | Average probability means ( $\hat{\mu}$ ) | Standard deviation ($\sigma$) | Average uncertainty index ($U = \sigma/\mu$) |
|---|---|---|---|---|---|
| A1 | Structural highs, clear reflections | 1500 | 0.87 | 0.08 | 0.092 |
| A2 | Anticline flanks, good continuity | 1300 | 0.81 | 0.11 | 0.136 |
| B1 | Fault-intensive areas, high SNR | 1200 | 0.76 | 0.15 | 0.197 |
| B2 | Regions with strong lithologic variations | 1000 | 0.65 | 0.18 | 0.277 |
| C1 | High noise interference | 900 | 0.54 | 0.21 | 0.389 |
| C2 | Thin interbedded sedimentary environments | 800 | 0.49 | 0.24 | 0.490 |
| D1 | Deep areas with low SNR | 700 | 0.42 | 0.25 | 0.595 |
| D2 | Unlabeled or blind areas | 600 | 0.37 | 0.28 | 0.757 |

Abbreviation: SNR: Signal-to-noise ratio.

similarity and curvature), fully supervised convolutional neural network methods, a semi-supervised framework, and the proposed self-supervised pre-training + transfer learning method. We also conducted ablation experiments, gradually removing modules such as mask reconstruction, 3D contrastive learning, and multimodal contrast to examine the contribution of each component.

In terms of evaluation metrics, we defined fault segmentation accuracy, recall, F1-score, and IoU to measure segmentation performance. Furthermore, to better meet the needs of seismic interpretation, we introduced the consistency index (CI) to measure the consistency between model predictions and those of different geological interpreters. Let the set of ground truth faults be $Y$ and the set of predictions be $\hat{Y}$, then the IoU is defined as:

$$\mathrm{IoU} = \frac{\left|Y \cap \hat{Y}\right|}{\left|Y \cup \hat{Y}\right|} \tag{17}$$

The consistency index CI is defined as the weighted ratio of the intersection and union between multiple interpreters:

$$\mathrm{CI} = \frac{1}{M}\sum_{i=1}^{M} \frac{\left|Y_i \cap \hat{Y}\right|}{\left|Y_i \cup \hat{Y}\right|} \tag{18}$$

where $Y_i$ is the fault label of the $i$-th interpreter, and $M$ is the number of interpreters.

Table 6 systematically compares the performance of five different methods in the fault identification task, clearly showing their respective technical characteristics and effect differences.

Beyond in-domain evaluation, we further validated the model's generalization capability on unseen data through a cross-dataset transfer experiment. Specifically, we performed self-supervised pre-training and supervised fine-tuning solely on the F3 dataset from the Dutch North Sea. The trained model was then applied directly for inference on the SEAM Phase I deepwater salt basin synthetic dataset, without any re-training or fine-tuning using SEAM labels. This setup simulates the model's performance when faced with a completely new geological environment (transitioning from a relatively simple deltaic setting to a complex salt-influenced regime) and different acquisition parameters.

Table 7 presents the fault segmentation results of our model on the SEAM dataset, comparing it against a supervised model trained directly on SEAM (Oracle, serving as an upper-performance bound) and a supervised model trained on F3 (without pre-training).

The results demonstrate that when a standard supervised model trained on F3 is directly transferred to SEAM, its performance drops sharply, achieving an IoU of only 0.42 on SEAM compared to 0.72 on its native F3 test set. This significant performance degradation indicates substantial overfitting to the geological context of the

**Table 6. Performance comparison of different methods in fault identification tasks**

| Method category | Accuracy | Recall | F1-score | IoU | CI |
|---|---|---|---|---|---|
| Similarity attribute + threshold detection | 0.74 | 0.58 | 0.65 | 0.49 | 0.61 |
| Curvature attribute method | 0.77 | 0.61 | 0.68 | 0.52 | 0.64 |
| Supervised convolutional neural network | 0.88 | 0.79 | 0.83 | 0.72 | 0.81 |
| Semi-supervised method | 0.85 | 0.75 | 0.79 | 0.68 | 0.78 |
| Self-supervised + transfer learning in this paper | 0.91 | 0.83 | 0.87 | 0.76 | 0.85 |

Abbreviations: CI: Consistency index; IoU: Intersection of union.

**Table 7. Cross-dataset generalization results (trained on F3 → tested on SEAM)**

| Method | Precision | Recall | F1-score | IoU |
|---|---|---|---|---|
| Supervised model (trained on SEAM, Oracle) | 0.89 | 0.81 | 0.85 | 0.74 |
| Supervised model (trained on F3, no pre-training) | 0.68 | 0.52 | 0.59 | 0.42 |
| GeoSSL (pre-trained & fine-tuned on F3, tested on SEAM) | 0.82 | 0.74 | 0.78 | 0.65 |

Abbreviation: IoU: Intersection of union.

training data. In contrast, our GeoSSL model, leveraging its robust feature representations learned through multi-task self-supervision on F3, exhibits superior generalization, achieving an F1-score of 0.78 and an IoU of 0.65 on SEAM. While there is still a gap compared to the Oracle model trained directly on SEAM (IoU = 0.74), the performance degradation of the GeoSSL model is far less severe than that of the purely supervised model. This strongly suggests that the GeoSSL framework learns more universal and transferable geological features, rather than merely memorizing the texture of a specific dataset. This result directly addresses the challenge of geological variability in practical applications and validates the model's significant potential for real-world exploration scenarios.

Looking at overall performance trends, traditional seismic attribute methods (similarity and curvature attributes) perform relatively poorly across all metrics. The similarity attribute plus threshold detection method exhibits the weakest overall performance, with an accuracy of 0.74, a recall of only 0.58, an F1-score of 0.65, an IoU of 0.49, and a confidence index of 0.61. This indicates that traditional methods suffer from significant underdetection issues in fault identification. The curvature attribute method shows slight improvement, with all metrics showing improvement, but still hovering between 0.6 and 0.7, reflecting the limitations of handcrafted feature-based methods. Supervised convolutional neural network methods demonstrate significant advantages, reaching an accuracy of 0.88, an F1-score of 0.83, an IoU of 0.72, and a confidence index of 0.81, demonstrating the powerful capabilities of deep learning for feature learning. Semi-supervised methods maintain good performance despite limited labeled data, achieving an accuracy of 0.85 and an F1-score of 0.79, but performing slightly worse than fully supervised methods. The self-supervised combined with transfer learning approach proposed in this paper achieved state-of-the-art results across all evaluation metrics, with precision reaching 0.91, recall reaching 0.83, F1-score of 0.87, IoU of 0.76, and confidence index of 0.85. This significant improvement in recall demonstrates that this approach effectively reduces missed detections and comprehensively outperforms traditional and supervised learning methods, demonstrating the superior performance and practical value of combining self-supervised pre-training with transfer learning in fault identification.

Figure 8 provides a visual comparison of the different approaches on a 3D seismic volume.

While the traditional seismic attribute method in Figure 8A can reflect the continuity and discontinuity of stratigraphic layers to a certain extent, the fault boundaries are blurred, making it difficult to form a coherent structural interpretation. The supervised neural network in Figure 8B can extract major faults when adequately labeled but still exhibits incomplete and spurious responses in locations with significant fault strike variations. Figure 8C shows the fault probability volume output by the SSL method proposed in this paper. Its advantages are clear, with clear fault boundaries and significantly improved connectivity. The model is particularly able to maintain the geometric integrity of faults in complex intersections. This demonstrates that self-supervised pre-training helps the model learn more robust multi-scale features, enabling it to achieve performance comparable to or even exceeding that of supervised methods even with limited annotation.

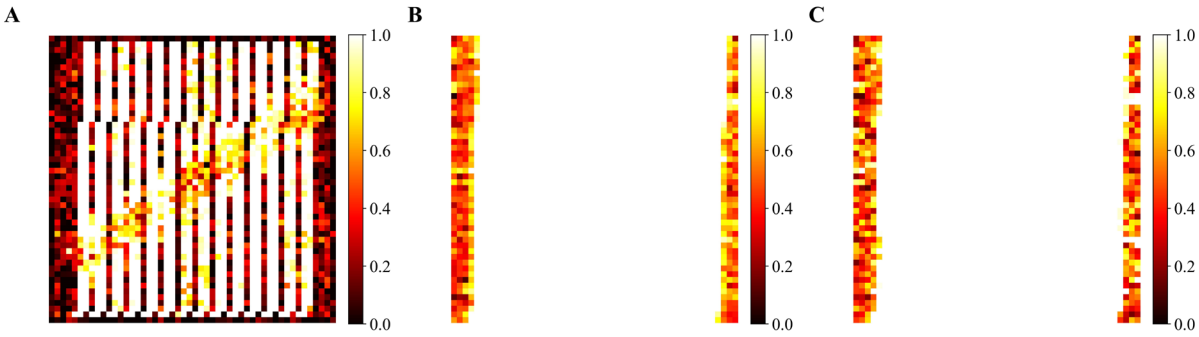



**Figure 8.** Comparative analysis of fault detection methods through 3D seismic data visualization. (A) Traditional attribute-based fault detection showing gradient-magnitude approach with significant noise artifacts and limited fault definition; (B) Supervised CNN fault identification demonstrating improved edge detection with moderate noise reduction and fault boundary enhancement; (C) Proposed self-supervised method displaying superior fault delineation with multi-scale feature fusion, minimal background noise, and enhanced detection of weak fault systems.
Abbreviation: CNN: Convolutional Neural Network.

To validate the contribution of each module, we conducted ablation experiments. Table 8 shows the performance of the model on the test set as we gradually remove different self-supervised tasks.

Removing the multimodal contrast module resulted in a significant drop in all metrics, with accuracy dropping to 0.89, F1-score to 0.85, and IoU to 0.73. This demonstrates the importance of multimodal information fusion in improving model performance. Removing the 3D contrast learning module further degraded performance, with accuracy dropping to 0.87, recalling to 0.78, and F1-score to 0.82. This demonstrates the critical value of 3D spatial feature learning in fault identification. Removing the mask reconstruction module resulted in an even more significant performance drop, with accuracy dropping to only 0.86, recall to 0.76, and F1-score to 0.80, indicating that the mask reconstruction pre-training strategy significantly impacted feature learning. Most notably, the purely supervised fine-tuning approach without self-supervised pre-training achieved the worst performance, with a precision of only 0.84, a recall of 0.73, an F1-score of 0.78, an IoU of 0.66, and a confidence index of 0.76. This is a significant gap compared to the full model, demonstrating the fundamental role and necessity of the self-supervised pre-training stage for subsequent supervised fine-tuning.

Figure 9 shows the comparative results for a typical profile, using human interpretation as a reference baseline.

As can be seen, the manual annotation in Figure 9A accurately outlines the main locations and strikes of the faults. While the semi-supervised method in Figure 9B generally agrees with the manual interpretation, the fault boundaries are generally rough, with some faults

**Table 8. Comparison of ablation experiment results**

| Model architecture | Accuracy | Recall | F1-score | IoU | CI |
|---|---|---|---|---|---|
| All modules (full model) | 0.91 | 0.83 | 0.87 | 0.76 | 0.85 |
| Removed multimodal contrast | 0.89 | 0.81 | 0.85 | 0.73 | 0.82 |
| Removed 3D contrastive learning | 0.87 | 0.78 | 0.82 | 0.70 | 0.80 |
| Removed mask reconstruction | 0.86 | 0.76 | 0.80 | 0.68 | 0.78 |
| No self-supervised pre-training (only supervised fine-tuning) | 0.84 | 0.73 | 0.78 | 0.66 | 0.76 |

Abbreviation: CI: Consistency index; IoU: Intersection of union.

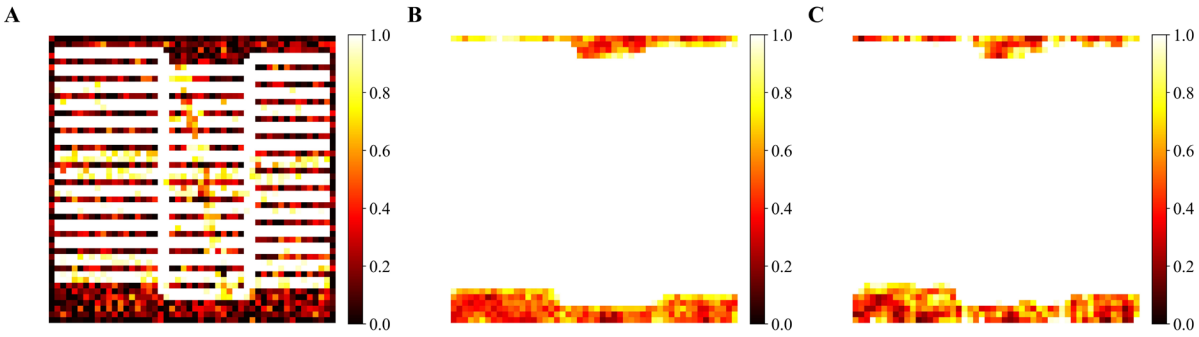



**Figure 9.** Cross-sectional comparison of fault interpretation methodologies on seismic profiles. (A) Manual expert interpretation showing traditional geological analysis with subjective fault boundary delineation based on geophysicist experience; (B) Semi-supervised learning approach demonstrating automated fault detection with moderate accuracy and some detail preservation; (C) Proposed self-supervised method exhibiting optimal fault identification with improved continuity, reduced false positives, and enhanced geological feature preservation.

appearing discontinuous or even obscured by background noise. In contrast, the self-supervised method in Figure 9C not only maintains a highly consistent fault distribution with the manual interpretation at large scales but also demonstrates greater coherence and refinement at small scales. In areas of localized weak reflection, the semi-supervised method struggles to capture small faults, while our proposed method effectively distinguishes signal from noise through pre-trained features, highlighting its greater tectonic sensitivity.

In a more in-depth analysis of the results, we further analyzed the model's performance under different geological conditions. Table 9 compares the performance in shallow areas with high SNRs, deep areas with low SNRs, areas with dense faults, and areas with weak faults.

Looking at overall performance trends, the model performs best in shallow, high-SNR areas, achieving accuracy of 0.93, recall of 0.86, F1-score of 0.89, IoU of 0.79, and a confidence index of 0.87. This is due to the high-quality seismic data providing the model with clear and reliable feature information. In deeper, low-SNR areas, due to seismic signal attenuation and increased noise interference, the model performance declines modestly, with accuracy of 0.88, recall of 0.79, F1-score of 0.83, IoU of 0.72, and a confidence index of 0.81. However, the model still maintains good recognition capabilities, indicating strong resistance to noise interference. In areas with dense faults, the model performs exceptionally well, achieving accuracy of 0.90, F1-score of 0.86, IoU of 0.75, and a confidence index of 0.84, demonstrating its ability to effectively handle the intersecting and interfering nature of complex fault systems. In the weak fault area, although the various indicators have declined (precision 0.87, recall

**Table 9. Model performance under different geological conditions**

| Geological conditions | Accuracy | Recall | F1-score | IoU | CI |
|---|---|---|---|---|---|
| Shallow high-SNR area | 0.93 | 0.86 | 0.89 | 0.79 | 0.87 |
| Deep low-SNR area | 0.88 | 0.79 | 0.83 | 0.72 | 0.81 |
| Fault-intensive area | 0.90 | 0.82 | 0.86 | 0.75 | 0.84 |
| Weak fault area | 0.87 | 0.80 | 0.83 | 0.71 | 0.82 |

Abbreviation: CI: Consistency index; IoU: Intersection of union; SNR: Signal-to-noise ratio.

0.80, F1-score 0.83, IoU 0.71, confidence index 0.82), the overall performance is still satisfactory, proving that the model has good sensitivity to the detection of weak fault signals.

Figure 10 focuses on the comparison of the identification effect of weak fault areas. This type of fault is usually easily overlooked due to its small amplitude difference and low SNR.

In the original seismic section displayed in Figure 10A, the fault signal is almost obscured by background reflections, making it difficult to identify directly with the naked eye. While the supervised model in Figure 10B enhances the response to anomaly locations to some extent, the results remain fragmented, with the faults barely connected. Figure 10C shows the predictions from our proposed method, where the fault zone is clearly highlighted, with continuous boundaries that align with the actual geological features. These results demonstrate the unique advantages of SSL in capturing subtle structural information and effectively improving the ability to identify low-energy faults, which is particularly critical for fault interpretation in deep, low-signal-to-noise environments.

To provide a more rigorous assessment of weak fault detection performance, we present a direct visual comparison between our GeoSSL predictions and manual expert interpretation in a representative low-SNR region, as shown in Figure 11. This region, located in the deeper section of the real-field dataset, is characterized by significant seismic energy attenuation and complex noise patterns that mask subtle fault displacements.

In Figure 11A, the original seismic section exhibits poor reflector continuity and pervasive background noise, making manual fault identification challenging even for experienced interpreters. The expert interpretation in Figure 11B identifies the major fault zone in the central portion of the section but shows hesitation in delineating its precise boundaries and omits several subtle fault splays visible only as minor disruptions in reflector alignment. Figure 11C demonstrates that our GeoSSL model successfully highlights the main fault zone with well-defined boundaries and, notably, detects multiple secondary fault splays that were missed by the manual interpreter. The overlay in Figure 11D reveals strong spatial agreement between manual and automated interpretations in the core fault zone (purple regions), while the model's predictions extend fault continuity into areas where the interpreter could not confidently assign labels due to noise interference.

Despite this strong performance, the comparison also reveals inherent limitations of our method in weak-signal areas. In the lower-left corner of Figure 11C, the model generates a low-probability response (light blue) that does not correspond to any interpreted fault in Figure 11B. Manual re-examination of this location suggests it may represent either a subtle stratigraphic feature misidentified as a fault or a genuine sub-seismic fracture below the resolution of manual interpretation. This ambiguity highlights a fundamental challenge in weak-signal detection: the absence of ground truth labels in noisy regions makes it difficult to distinguish between true positive detections and false positives arising from noise artifacts.

Combining these experimental results, we conclude that self-supervised pre-training significantly improves the model's generalization and robustness, enabling it to produce high-quality fault imaging results even with limited annotations. Ablation experiments confirm

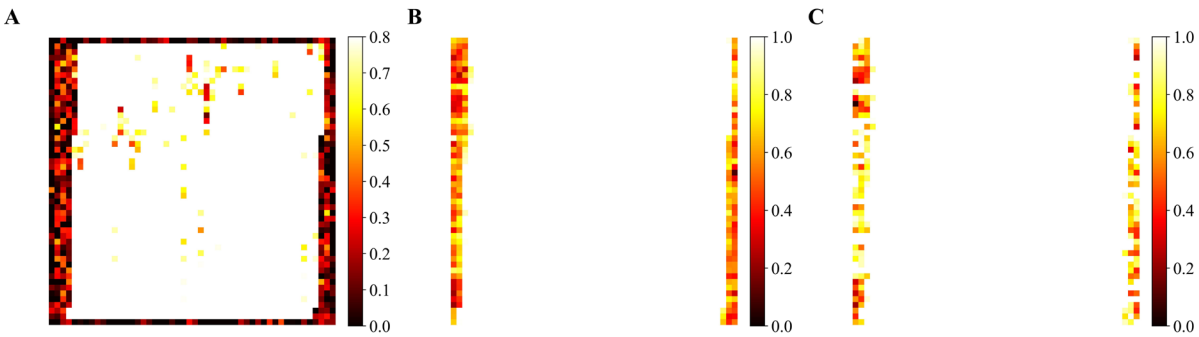



**Figure 10.** Weak fault identification performance comparison across different computational approaches. (A) Original seismic section containing subtle fault displacements and low-amplitude fault-related features challenging for conventional detection methods; (B) Supervised CNN results showing partial identification of prominent faults but limited capability for weak fault detection; (C) Proposed self-supervised approach demonstrating enhanced sensitivity to weak fault signals with improved SNR and comprehensive fault network delineation.
Abbreviation: CNN: Convolutional Neural Network; SNR: Signal-to-noise ratio.

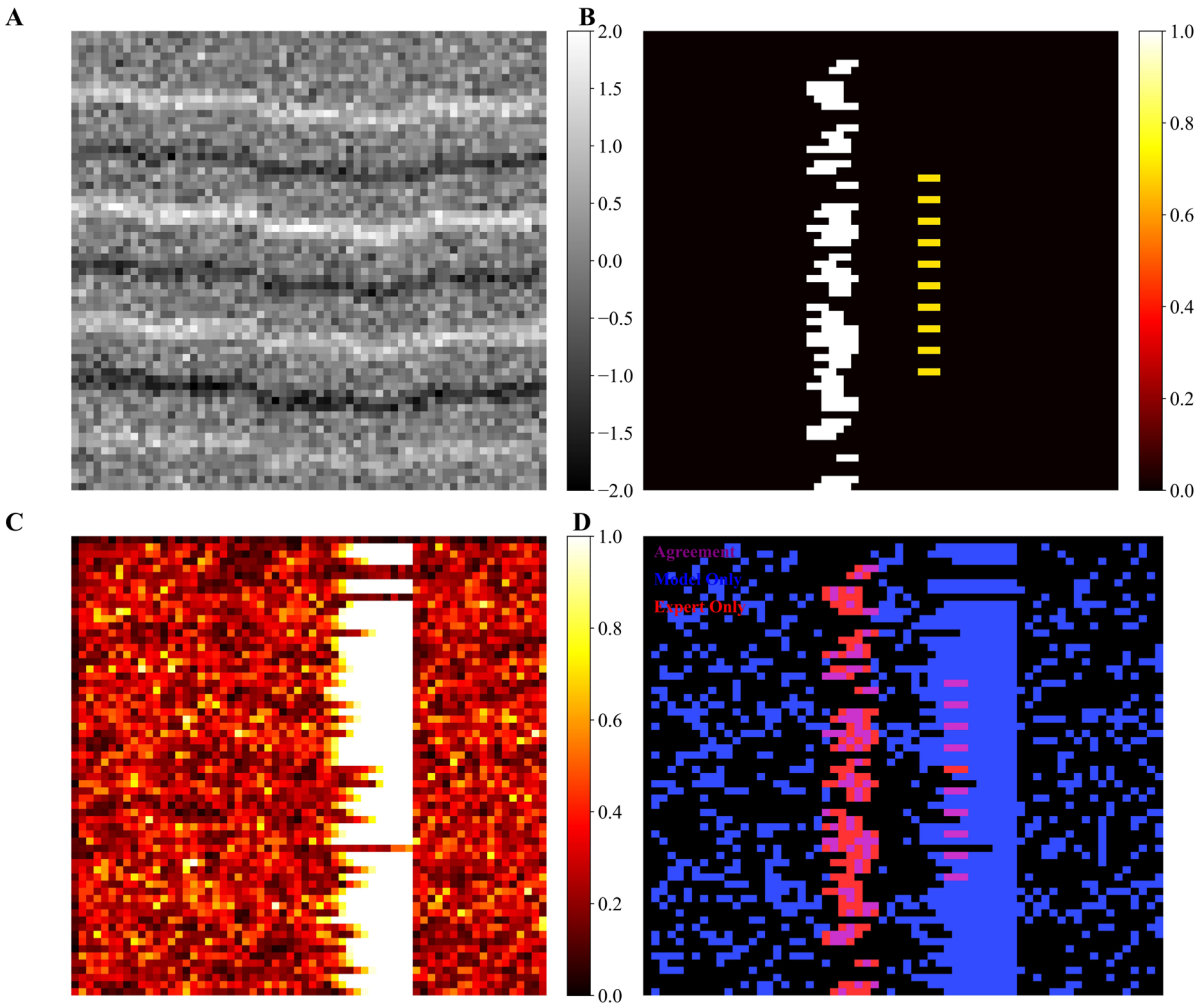



**Figure 11.** Comparative visualization of weak fault detection in a low signal-to-noise ratio region. (A) Original seismic section with weak fault signals obscured by noise; (B) Manual expert interpretation highlighting the interpreted fault traces; (C) GeoSSL-predicted fault probability map; (D) Overlay of manual interpretation (red) and GeoSSL predictions (blue) with agreement shown in purple.

that mask reconstruction, contrastive learning, and multimodal features all contribute significantly to the performance improvements. Our proposed method performs particularly well in weak signal areas, providing a more reliable technical tool for subsequent geological interpretation and exploration risk control.

## 7. Model generalization and practical application verification

After constructing an SSL-based fault identification framework, we further focused on its generalization capabilities across different seismic data types and actual work areas. The model's generalization is primarily reflected in its adaptability across seismic sources, frequency bands, and tectonic settings. For a systematic evaluation, we conducted migration tests on multiple sets of seismic data, including data from multi-channel source acquisition systems, filtered data in different frequency bands, and field data from different sedimentary environments.

The model's transferability was quantified using the following metrics: the IoU between the fault probability volume and the manually annotated data, the CI, and the stability of the predictions on new data ( $S$ ).[34] Stability is defined as:

$$S = 1 - \frac{\sigma_P}{\mu_P} \tag{19}$$

where $\mu_P$ represents the mean fault probability predicted by the model under multiple random augmentations, and $\sigma_P$ represents the standard deviation of the prediction. This metric reflects the robustness of the model under different data augmentation or noise conditions.

Figure 12 shows the generalization effect of the model under different earthquake source and frequency band conditions.

Figure 12A shows the fault probability volume for broadband data, Figure 12B presents the low-frequency filtering result, and Figure 12C depicts the high-frequency enhanced data result. According to these results, even with large variations in frequency band distribution, the model is able to reliably identify fault boundaries, demonstrating the robustness of its self-supervised features to spectral variations.

To further analyze the model's performance in a real-world work area, we selected a structurally complex block and invited three experienced interpretation experts to perform manual annotations. The results were then compared with the model's predictions. Table 10 summarizes the consistency metrics between the model's predictions and the expert interpretations.

From the inter-expert consistency analysis, there are certain differences in the interpretation results among the three experts, but the overall consistency is high. The consistency between Experts 1 and 2 is the best (IoU = 0.81, F1-score = 0.87, CI = 0.89), and the consistency between Experts 1 and 3 is relatively low (IoU = 0.77, F1-score = 0.84, CI = 0.86). The average inter-expert consistency reaches IoU = 0.79, F1-score = 0.85, CI = 0.88, reflecting the inherent subjective differences in seismic interpretation. Comparisons between the model and the experts show that the model has the highest agreement with Expert 2 (IoU = 0.76, F1-score = 0.83, CI = 0.84),

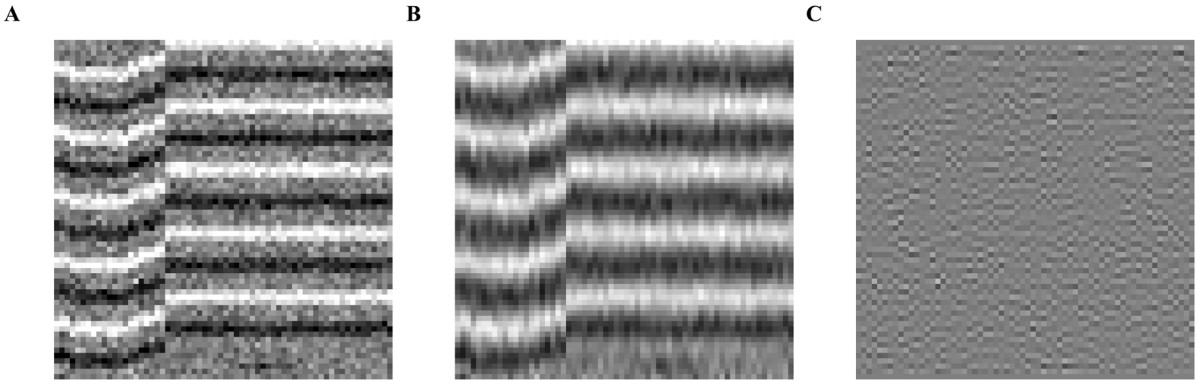



**Figure 12.** Frequency-domain analysis of fault detection: broadband versus filtered seismic data comparison. (A) Broadband seismic data preserving complete frequency spectrum with all reflector layers and fault features visible across multiple scales; (B) Low-frequency filtered data (0–20Hz) emphasizing major structural interfaces and large-scale geological features with reduced high-frequency noise; (C) High-frequency filtered data (30–50Hz) highlighting fine-scale fault discontinuities and detailed geological boundaries with enhanced fault edge definition.

**Table 10. Comparison of consistency between model predictions and expert interpretation results**

| Comparison mode | IoU | F1-score | CI | Stability |
|---|---|---|---|---|
| Expert 1 vs. Expert 2 | 0.81 | 0.87 | 0.89 | - |
| Expert 2 vs. Expert 3 | 0.79 | 0.85 | 0.88 | - |
| Expert 1 vs. Expert 3 | 0.77 | 0.84 | 0.86 | - |
| Model vs. Expert 1 | 0.74 | 0.82 | 0.83 | 0.91 |
| Model vs. Expert 2 | 0.76 | 0.83 | 0.84 | 0.92 |
| Model vs. Expert 3 | 0.73 | 0.81 | 0.85 | 0.90 |
| Model vs. Expert average | 0.74 | 0.82 | 0.84 | 0.91 |
| Average across experts | 0.79 | 0.85 | 0.88 | - |

Abbreviations: CI: Consistency index; IoU: Intersection of union.

while agreement with Expert 3 is relatively low (IoU = 0.73, F1-score = 0.81). The average agreement between the model and the experts is IoU = 0.74, F1-score = 0.82, CI = 0.84. While slightly lower than the inter-expert agreement, it is still quite close. Particularly noteworthy is the excellent performance of the model in terms of stability metrics, with stability against the three experts reaching 0.91, 0.92, and 0.90, respectively, with an average of 0.91. This result significantly outperforms the human experts (inter-expert stability cannot be quantitatively compared), demonstrating that the model has highly consistent and repeatable interpretation performance.

Figure 13 shows a comparison of a typical section in the work area. Figure 13A shows the expert interpretation results, Figure 13B presents the fault probability volume predicted by our method, and Figure 13C depicts the difference between the two.

According to the results, the model predictions and manual interpretations highly overlap for the locations of major faults. However, in areas with blurred boundaries, the model's probability values are lower, reflecting the uncertainty of the predictions and providing an intuitive confidence level for interpreters.

At the engineering application level, we further evaluated the model's practical value in reservoir trap evaluation and fault seal analysis. Fault seal is typically characterized by the shale gouge ratio (SGR) and the ratio of fault displacement to thickness. We combined the model's predictions of fault plane locations to calculate the following seal parameters:

$$SGR = \frac{\sum_{i=1}^{n} V_{sh,i} \cdot \Delta h_i}{D} \tag{20}$$

where $V_{sh,i}$ represents the shale content of layer $i$, $\Delta h_i$ represents its thickness, and $D$ represents the fault displacement.

Table 11 presents the results of the seal evaluation for a reservoir block. In areas with high SGR, the model-predicted fault probability volume is highly consistent with the oil and gas shows, validating its potential for application in oil and gas trap evaluation.

The seal evaluation results in Table 11 demonstrate strong consistency with geological principles and provide actionable insights for exploration decision-making. From a geological perspective, the SGR values correlate directly with the lithological context of the fault zones. Faults F2 and F7, both rated as "excellent" seals (SGR ≥ 0.80), are located in areas dominated by thick shale sequences, as confirmed by well logs in the adjacent blocks. The high mud ratio (≥38%) in these fault zones promotes the formation of continuous shale gouge along the fault plane, effectively creating a lateral seal for adjacent reservoir compartments. Conversely, Fault F5, rated as "poor" (SGR = 0.48), transects a sand-rich channel complex where the low mud ratio (22%) inhibits the development of an effective seal, making it a potential leakage pathway.

The integration of model-derived fault probability with SGR analysis offers significant practical value for exploration risk assessment. For instance, the combination of high fault detection confidence (F2: 0.83, F7: 0.81) and excellent seal capacity identifies these faults as critical structural elements for trap definition. In a prospect ranking scenario, traps bounded by F2 or F7 would be prioritized for drilling, as they present a lower risk of hydrocarbon leakage. In contrast, traps relying on F5 for closure would require careful evaluation of alternative

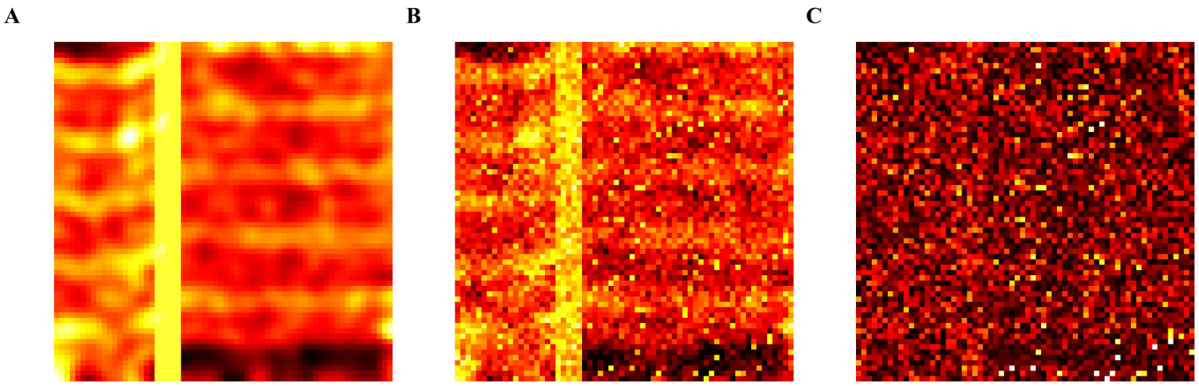



**Figure 13.** Field data validation: expert geological interpretation versus automated model predictions. (A) Expert geological interpretation based on traditional seismic analysis showing manually delineated fault boundaries with geophysicist experience and subjective geological reasoning; (B) Automated model predictions using the proposed self-supervised approach displaying computed fault probabilities with quantitative confidence measures; (C) Difference map highlighting discrepancies between expert interpretation and model predictions, revealing areas of agreement and potential interpretation uncertainty.

**Table 11. Model prediction results of faults sealing and reservoir trap evaluation**

| Fault number | Average fault probability | Displacement (m) | Mud ratio (%) | SGR | Trap integrity assessment |
|---|---|---|---|---|---|
| F1 | 0.86 | 45 | 32 | 0.72 | Good |
| F2 | 0.83 | 38 | 40 | 0.81 | Excellent |
| F3 | 0.68 | 27 | 25 | 0.55 | Fair |
| F4 | 0.72 | 34 | 29 | 0.61 | Good |
| F5 | 0.64 | 20 | 22 | 0.48 | Poor |
| F6 | 0.79 | 42 | 35 | 0.77 | Good |
| F7 | 0.81 | 36 | 38 | 0.80 | Excellent |
| F8 | 0.69 | 25 | 28 | 0.58 | Fair |

Abbreviation: SGR: Shale gouge ratio.

sealing mechanisms or be considered higher-risk prospects. Furthermore, the moderate SGR values of F1 (0.72) and F6 (0.77), combined with their high displacement (≥42 m), suggest they could function as effective baffles, potentially leading to compartmentalization within a reservoir. This information is crucial for field development planning, as it influences well placement and production strategy to account for potential flow barriers.

The consistency between model predictions and independent geological evidence—such as the alignment of high-SGR faults with mapped shale-rich intervals and the correspondence of low-SGR faults with sand-prone fairways—further validates the reliability of our GeoSSL-derived fault interpretations. By translating abstract fault probability volumes into concrete, decision-relevant parameters (seal capacity, compartmentalization risk), our framework bridges the gap between automated seismic interpretation and practical exploration workflows.

Figure 14 further illustrates the overlay of the fault probability volume and the sedimentary facies distribution.

Based on the findings, high-probability fault planes closely match sedimentary sandbody boundaries, providing a reliable basis for predicting reservoir connectivity and oil and gas migration directions. This demonstrates that our method is not only applicable to fault geometry research but also has applications in engineering, particularly risk control and deployment optimization in exploration and development.

In summary, the experimental results demonstrate that the proposed self-supervised fault identification method has good generalization capabilities. It not only maintains high performance under different seismic acquisition conditions and data preprocessing strategies but also achieves high consistency with expert interpretations in actual work areas. Combined with closure evaluation and sedimentary environment interpretation, the model's predictions provide solid technical support for reservoir trap analysis and oil and gas exploration decision-making.

## 8. Discussion

This study proposes an innovative SSL-based approach for 3D seismic fault identification, demonstrating several breakthroughs compared to existing techniques. First, we design a multi-task pre-training framework that integrates mask reconstruction, 3D patches contrast learning, and joint comparison of multimodal attributes. This integrated approach fully leverages unlabeled seismic data to learn geologically meaningful deep feature representations, significantly reducing reliance on expensive manual annotation. Second, we develop a boundary-sensitive module and multi-scale feature extraction mechanism.

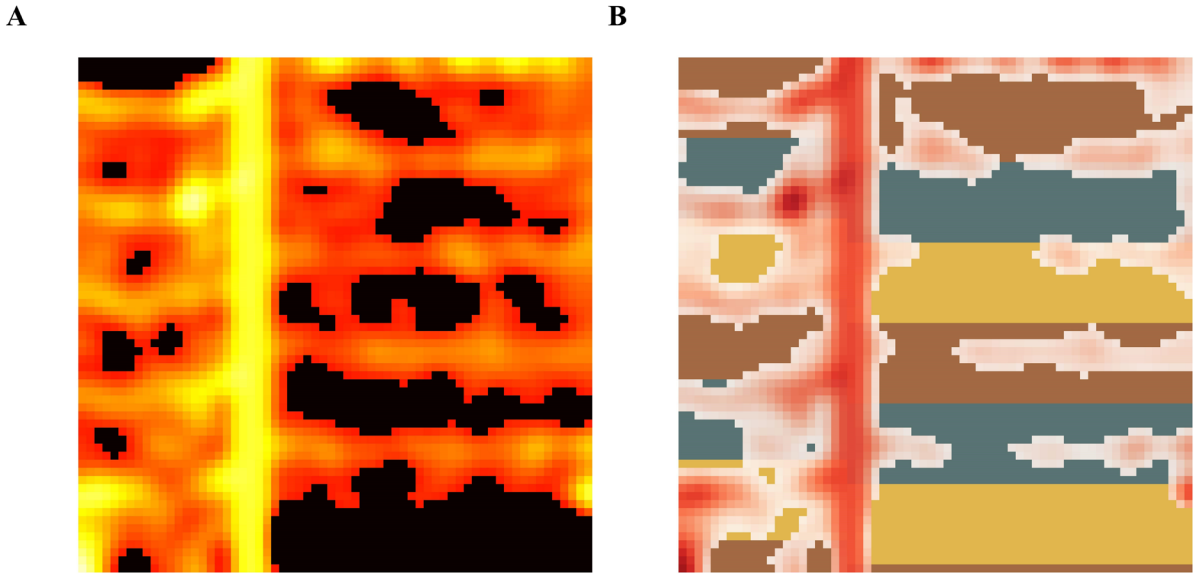



**Figure 14.** Integrated structural-stratigraphic analysis: fault systems overlaid on sedimentary facies distribution. (A) Fault probability volume showing high-confidence structural discontinuities with filtered background noise for clear fault delineation; (B) Comprehensive geological interpretation combining fault network overlay on sedimentary facies map with layered stratigraphy including sandstone channels, shale intervals, and mixed lithofacies, demonstrating the spatial correlation between structural deformation and depositional environments.

Through gradient weight enhancement and structural tensor analysis, we effectively enhance the model's ability to perceive fault edge features, particularly in weakly reflective regions. Third, we achieve smooth transfer from self-supervised pre-training to downstream tasks, achieving excellent performance across multiple tasks, including fault segmentation, classification, and strike and dip prediction, demonstrating the method's versatility and practicality. Finally, we introduce uncertainty quantification metrics and a stability assessment system to provide a reliable reference for fault interpretation results, a crucial capability lacking in traditional methods.

Self-supervised learning demonstrates significant potential for widespread application in the field of seismic fault identification. This study demonstrates that SSL technology can effectively address the scarcity of annotated data in seismic interpretation, providing a viable solution for the intelligent interpretation of large-scale 3D seismic data. With the continuous advancement of seismic acquisition technology and the continued growth of data volumes, SSL methods will become an important technical approach in the field of seismic interpretation.

Although this study has achieved significant results, several limitations remain that require further improvement. First, the current model is primarily based on seismic amplitude data and does not fully integrate multi-source geophysical data, such as gravity, magnetoelectric, and wellbore data. This limits the model's comprehensive interpretation capabilities in complex geological conditions. Future research should consider multi-source data fusion strategies and improve the model's comprehensive interpretation performance through cross-modal learning. Second, the model lacks clear physical constraints, and the learning process is primarily data-driven, which can lead to geologically unreasonable predictions. Introducing prior knowledge, such as seismic wave propagation equations and rock physics relationships, and developing a physically constrained SSL framework will be important areas for improvement. Third, the model's adaptability to extreme geological conditions (such as subsalt structures and volcanic rock cover) remains to be verified, requiring further testing and optimization in these challenging scenarios.

While our framework achieves high accuracy, its computational efficiency for production-scale deployments warrants consideration. On our hardware platform, processing a typical 400 × 400 × 1000 voxel volume requires approximately 73 s, with a GPU memory footprint of 3.2 GB per inference patch. For extremely large surveys exceeding billions of voxels, this translates to several minutes of computation on multi-GPU systems. Future work will explore optimization strategies such as model quantization, knowledge distillation, and sparse inference to further reduce computational overhead while maintaining detection accuracy, enabling seamless integration into time-sensitive exploration workflows.

Future research will focus on addressing these limitations and, through multidisciplinary integration, promote the development of 3D seismic fault identification technology toward higher accuracy and stronger generalization capabilities. Furthermore, while our current model focuses on structural discontinuities at the seismic scale, understanding fault properties fundamentally relates to the physics of interfaces—a subject of intense study across multiple scales and disciplines. Cross-disciplinary insights into interface physics could potentially inspire new geologically informed constraints or multi-scale modeling approaches for future generations of fault interpretation algorithms.

Beyond fault detection, the proposed GeoSSL framework holds significant potential for generalization to other fundamental seismic interpretation tasks. The core innovation—learning transferable, geology-aware representations from unlabeled data through multi-task self-supervision—is not inherently task-specific. The same pre-trained encoder, which captures both local textural patterns and global structural context, could be readily adapted to related tasks with minimal modification. For instance, in horizon picking, the model's sensitivity to lateral continuity and reflector termination could be leveraged to automate the tracking of stratigraphic surfaces. In channel detection, the framework's ability to recognize complex geometric patterns and amplitude anomalies could facilitate the delineation of fluvial or turbidite systems. Furthermore, the multimodal contrastive learning mechanism, designed to fuse complementary attribute information, could prove particularly valuable for lithofacies classification, where integrating amplitude, phase, and frequency attributes is crucial. Exploring these extensions represents a promising direction for developing a unified, foundation-model-style approach to seismic interpretation, where a single pre-trained model serves as a versatile backbone for multiple downstream geological analysis tasks.

## 9. Conclusion

In conclusion, this study presents an innovative methodology for 3D seismic fault identification by pioneering a dedicated SSL framework for geophysical data. The primary innovation is the synergistic integration of complementary pre-training tasks, which for the first time enables a model to autonomously learn comprehensive, transferable representations of fault systems from vast unlabeled volumes. This is further enhanced by our novel boundary-aware network design and demonstrated effectiveness across a suite of interpretation tasks with minimal supervision. Our work transcends incremental improvement of existing models; it provides a blueprint for building trustworthy, geology-aware AI interpreters, significantly reducing the annotation bottleneck and setting a new direction for intelligent geoscience analysis.

This research has positive implications for the advancement of the field of fault identification and interpretation in 3D seismic data. First, the SSL paradigm successfully addresses the bottleneck of limited labeled samples, a major constraint on the practical application of deep learning. This enables automated and refined interpretation of large-scale 3D seismic volumes. Second, this method exhibits strong cross-data generalization capabilities, adapting to seismic data with varying acquisition parameters, frequency band characteristics, and geological backgrounds, enhancing the technology's practicality and stability in real-world exploration scenarios. More importantly, the effective integration of fault identification results with geological information such as sedimentary facies and closure parameters provides a more reliable structural basis for reservoir evaluation, oil and gas detection, and drilling deployment, helping to reduce exploration and development risks and improve economic efficiency.

In subsequent research, we will further refine and improve self-supervised fault identification technology from multiple perspectives. First, we will explore multimodal data fusion mechanisms, integrating multiple sources of information, such as gravity, magnetoelectricity, and well data, to enhance the model's comprehensive reasoning capabilities under complex geological conditions. Second, we will introduce physics-driven constraints, embedding prior knowledge, such as seismic wave propagation equations and rock physics relationships, into the SSL process to enhance the geological plausibility of the model output. We will also optimize the model's structure and computational efficiency to better suit the needs of large-scale 3D seismic data processing and real-time interpretation and explore extending its application to more challenging geological environments such as subsalt structures and volcanic areas. Ultimately, through continuous technological innovation and practical verification, we aim to advance seismic interpretation toward a higher level of intelligence and automation, providing stronger technical support for fields such as energy exploration and geological hazard prediction.

## Acknowledgments

None.

## Funding

This project was supported by the Hubei Provincial Engineering Technology Research Center for Information Technology in Clothing Industry Open Fund (Project No.: 2025HBCI03), the Hanjiang Normal University Scientific

Research Fund Project (2024B32), and the CCF-Zhipu Large Model Fund Project (202212).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Mengge Wang, Junyu Zhang, Haofa Lin
*Formal analysis:* Mengge Wang, Xinrong Hu
*Investigation:* Junyu Zhang, Haofa Lin
*Methodology:* Xinrong Hu, Junyu Zhang
*Validation:* Mengge Wang, Xinrong Hu
*Writing–original draft:* Mengge Wang, Xinrong Hu, Junyu Zhang
*Writing–review & editing:* Mengge Wang, Xinrong Hu, Junyu Zhang, Haofa Lin

## Availability of data

The 3D seismic data used in this study were obtained from the publicly available F3 Block dataset, provided by TNO and dGB Earth Sciences via the OpendTect platform. The authors are grateful for the unrestricted data access.

## References

1. Alao JO. The emerging roles of 3D and 4D geophysical and geological modelling in evaluating seismic risks: A critical review. *Earthquake Res Adv*. 2025;6(1):100399.

   doi: 10.1016/j.eqrea.2025.100399

2. Seyyedattar M, Zendehboudi S, Butt S. Technical and non-technical challenges of development of offshore petroleum reservoirs: Characterization and production. *Nat Resour Res*. 2020;29(3):2147-2189.

   doi: 10.1007/s11053-019-09549-7

3. Daramola GO, Jacks BS, Ajala OA, Akinoso AE. Enhancing oil and gas exploration efficiency through AI-driven seismic imaging and data analysis. *Eng Sci Technol J*. 2024;5(4):1473-1486.

   doi: 10.51594/estj.v5i4.1077

4. Verma S, Chopra S, Ha T, Li F. A review of some amplitude-based seismic geometric attributes and their applications. *Interpretation*. 2022;10(1):B1-B12.

   doi: 10.1190/INT-2021-0136.1

5. Gui J, Chen T, Zhang J, *et al*. A survey on self-supervised learning: algorithms, applications, and future trends. *IEEE Trans Pattern Anal Mach Intell*. 2024;46(12):9052-9071.

   doi: 10.1109/TPAMI.2024.3415112

6. Yang Y, Wang Z, Liu N, *et al*. Physically driven self-supervised learning and its applications in geophysical inversion. *IEEE Trans Geosci Remote Sens*. 2024;62:4503211.

   doi: 10.1109/TGRS.2024.3368016

7. Chai X, Yang T, Gu H, Tang G, Cao W, Wang Y. Geophysics-steered self-supervised learning for deconvolution. *Geophys J Int*. 2023;234(1):40-55.

   doi: 10.1093/gji/ggad015

8. Salazar JJ, Maldonado-Cruz E, Ochoa J, Pyrcz MJ. Self-supervised learning for seismic data: Enhancing model interpretability with seismic attributes. *IEEE Trans Geosci Remote Sens*. 2023;61:1-18.

   doi: 10.1109/TGRS.2023.3285820

9. Zhang Z, Chen R, Ma J. Improving seismic fault recognition with self-supervised pre-training: a study of 3D transformer-based with multi-scale decoding and fusion. *Remote Sens*. 2024;16(5):922.

   doi: 10.3390/rs16050922

10. Sheng H, Wu X, Si X, Li J, Zhang S, Duan X. Seismic foundation model: A next generation deep-learning model in geophysics. *Geophysics*. 2025;90(2):IM59-IM79.

    doi: 10.1190/geo2024-0262.1

11. Dou Y, Li K. 3D seismic fault detection via contrastive-reconstruction representation learning. *Expert Syst Appl*. 2024;249:123617.

    doi: 10.1016/j.eswa.2024.123617

12. Wang J, Ma S, Liu Y, Dong R. AttentionFaultFormer: An attention-enhanced 3D CNN & transformer model for seismic fault detection. *J Appl Geophys*. 2025;238:105707.

    doi: 10.1016/j.jappgeo.2025.105707

13. Chen R, Zhang Z, Ma J. Seismic Fault SAM: adapting SAM with lightweight modules and 2.5D strategy for fault detection. In: Yuan B, Ruan Q, Wei S, An G, eds. In: *Proceedings of the2024 IEEE 17th International Conference on Signal Processing (ICSP2024)*; October 28-31, 2024; Suzhou, China. IEEE Press; 2024:436-441.

    doi: 10.1109/ICSP62129.2024.10846297

14. Zu S, Zhao P, Ke C, Junxing C. ResACEUnet: an improved transformer Unet model for 3D seismic fault detection. *J Geophys Res Mach Learn Comput*. 2024;1(3):e2024JH000232.

    doi: 10.1029/2024JH000232

15. Khosro Anjom F, Vaccarino F, Socco LV. Machine learning for seismic exploration: where are we and how far are we from the holy grail? *Geophysics*. 2024;89(1):WA157-WA178.

    doi: 10.1190/geo2023-0129.1

16. Monteiro BA, Oliveira H, dos Santos JA. Self-supervised learning for seismic image segmentation from few-labeled samples. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

    doi: 10.1109/LGRS.2022.3193567

17. Choi W, Pyun S, Jou HT. Synthetic training data optimization for enhanced fault detection in seismic images. *Lithosphere*.

2025;2025(3):lithosphere_2024_240.

doi: 10.2113/2025/lithosphere_2024_240

18. Noor UA. Machine learning innovations in revolutionizing earthquake engineering: a review. *Arch Comput Methods Eng*. 2025;33(1):687-743.

doi: 10.1007/s11831-025-10320-w

19. Stucchi E, Mazzotti A, Ciuffi S. Seismic preprocessing and amplitude cross-calibration for a time-lapse amplitude study on seismic data from the Oseberg reservoir. *Geophys Prospect*. 2005;53(2):265-282.

doi: 10.1111/j.1365-2478.2004.00471.x

20. Li X, Li K, Xu Z, Huang Z, Dou Y. Fault-Seg-Net: A method for seismic fault segmentation based on multi-scale feature fusion with imbalanced classification. *Comput Geotech*. 2023;158:105412.

doi: 10.1016/j.compgeo.2023.105412

21. Wang W, Huang Q, You S, Yang C, Neumann U. Shape inpainting using 3D generative adversarial network and recurrent convolutional networks. In: *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV 2017)*; October 22-29, 2017; Venice, Italy. IEEE; 2017:2317-2325.

doi: 10.1109/ICCV.2017.252

22. Wang W, Shen S, Yuan Y. Forecasting short-term export volumes with hybrid models integrating SARIMA with attention-based LSTM. *Int Sci Tech Econ Res*. 2026;4(1):1-22.

doi: 10.71451/ISTAER2601

23. Ding Y, Chen G. Joint prediction model of reservoir parameters based on multimodal transformer graph neural operator physical constraint network. *Int Sci Tech Econ Res*. 2026;4(1):70-89.

doi: 10.71451/ISTAER2604

24. 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

25. Wang Z, Bovik AC, Sheikh HR. Structural similarity based image quality assessment. In: Wu HR, Rao KR, eds. *Digital Video Image Quality and Perceptual Coding*. CRC Press; 2017:225-242.

26. Zheng J, Tang Y, Huang A, Wu D. Hierarchical multivariate representation learning for face sketch recognition. *IEEE Trans Emerg Top Comput Intell*. 2024;8(2):2037-2049.

doi: 10.1109/TETCI.2024.3359090

27. Sun P, Zheng Y, Xu W, Li J, Yang J. Completing missing entities: Exploring consistency reasoning for remote sensing object detection. *IEEE Trans Image Process*. 2026;35:569-584.

doi: 10.1109/TIP.2025.3648164

28. Belmouhcine A, Pham MT, Lefèvre S. YOLO-G3CF: Gaussian contrastive cross-channel fusion for multimodal object detection. *IEEE Geosci Remote Sens Lett*. 2025;22:8002005.

doi: 10.1109/LGRS.2025.3564181

29. Si X, Wu X, Sheng H, Zhu J, Li Z. SeisCLIP: A seismology foundation model pre-trained by multimodal data for multipurpose seismic feature extraction. *IEEE Trans Geosci Remote Sens*. 2024;62:5903713.

doi: 10.1109/TGRS.2024.3354456

30. Contreras V, Stewart JP, Kishida T, *et al*. NGA-Sub source and path database. *Earthq Spectra*. 2022;38(2):799-840.

doi: 10.1177/87552930211065054

31. Sun X, Xu Q, Yin Z, *et al*. Prediction equations for pulse parameters based on physics-based ground motion simulation. *Appl Geophys*. 2025.

doi: 10.1007/s11770-025-1311-z

32. Wu S, Wang Y. Seismic image dip estimation by multiscale principal component analysis. *IEEE Trans Geosci Remote Sens*. 2022;61:5900410.

doi: 10.1109/TGRS.2022.3229332

33. Yang Y, Jin BB, Sun X, *et al*. Exact counting of subtrees with diameter no more than d in trees: a generating function approach. *Inf Comput*. 2025;307:105353.

doi: 10.1016/j.ic.2025.105353

34. Zhou G, Li T, Li K, Chu S, Zhu X. Fault-GSA: high generalization 3D fault detection method based on sparse annotations. *IEEE Trans Geosci Remote Sens*. 2025;63:5910317.

doi: 10.1109/TGRS.2025.3557022

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# A two-branch convolutional neural network for pre-stack elastic parameter inversion of coal-bearing gas reservoirs

**Fei Li[1,2], Mengbo Zhang[1,2], Qiang Liang[1], Xiaojie Cui[1,2], Na Ni[1], Qingzhou Zhang[2], and Yongheng Zhang[3]***

[1]Exploration and Development Research Institute, Changqing Oilfield Company, PetroChina, Xi'an, Shaanxi, China

[2]National Engineering Laboratory for Exploration and Development of Low-Permeability Oil and Gas Fields, College of Petroleum Engineering, Xi'an, Shaanxi, China

[3]College of Geophysics, Chengdu University of Technology, Chengdu, Sichuan, China

## Abstract

Accurate estimation of Poisson's ratio is essential for the characterization of coal-bearing gas reservoirs, particularly in thin-bed and low signal-to-noise ratio environments where conventional elastic impedance (EI) inversion suffers from wavelet interference, limited resolution, and reliance on linearization assumptions. To address these limitations, we develop a physics-guided two-branch convolutional neural network (TB-CNN) that directly predicts Poisson's ratio by jointly integrating EI-inverted P-wave velocity, S-wave velocity, and density with small-, medium-, and large-angle seismic stacks. The first branch provides geologically consistent, physics-informed background trends, while the second branch captures thin-bed-sensitive reflectivity features and amplitude tuning effects. The fused latent representation is explicitly regularized using empirical rock-physics relationships to ensure physical plausibility and enhanced generalization. Field validation on the 8# coal seam of the Ordos Basin demonstrates that the proposed TB-CNN improves vertical resolution, sharpens seam boundary delineation, and better preserves thickness variations compared with EI inversion and a single-branch CNN. Near-well comparisons show higher correlation with log-derived Poisson's ratio, while lateral slices reveal improved continuity and thin-layer detectability. These results confirm that combining physics-guided stability with data-driven resolution provides a robust and interpretable framework for Poisson's ratio inversion in thin coal seams and holds promise for broader applications in unconventional gas reservoir prediction.

***Keywords:*** Coalbed methane; Elastic parameter inversion; Pre-stack seismic data; Convolutional neural network

***Corresponding author:**
Yongheng Zhang
(zhangyhzjk@163.com)

**Citation:** Li F, Zhang M, Liang Q, *et al*. A two-branch convolutional neural network for pre-stack elastic parameter inversion of coal-bearing gas reservoirs. *J Seismic Explor.* 2026;35(2):025390076.
doi: 10.36922/JSE025390076

**Received:** September 22, 2025

**Revised:** November 10, 2025

**Accepted:** January 28, 2026

**Published online:** April 23, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Elastic parameters—especially Poisson's ratio (ν) and Young's modulus (E)—are fundamental for assessing the mechanical behavior, brittleness, and gas-bearing potential of coal reservoirs. In coalbed methane (CBM) development, reliable ν and

E volumes support *in situ* stress prediction, hydraulic-fracturing design, and dynamic reservoir simulation, and thus are crucial for identifying sweet spots and evaluating completion strategies in thinly bedded, laterally heterogeneous coal seams, such as those in the Ordos Basin, China.[1,2] However, deriving these parameters from seismic data remains challenging due to insufficient resolution and stability.

Conventional pre-stack inversion frameworks—grounded in amplitude versus offset (AVO)/amplitude variation with angle theory and linearized reflectivity equations—are well established.[3,4] Yet, they are often sensitive to noise, wavelet, and angle calibration, and acquisition footprint, and typically rely on smooth initial models and strong regularization. As a result, the inverted P-wave velocity (Vp), S-wave velocity (Vs), and density (ρ) volumes tend to be band-limited and over-smoothed, which limits thin-bed detection and blurs sharp elastic contrasts. These limitations are particularly restrictive in coal-bearing settings where the target seams are thin and embedded within a single reflection lobe, as is the case for the 8# coal seam in the northern Yishan Slope of the Ordos Basin.[5,6] In the standard workflow, ν is conventionally derived from the Vs/Vp ratio, which inevitably causes errors from intermediate elastic parameter inversions to propagate and degrade the final prediction accuracy. Hence, any bias from intermediate inversions propagates and degrades the final attribute used for brittleness and stress analysis.[7-10]

Recent advances in deep learning (DL) offer an alternative by directly learning nonlinear mappings between seismic observations and rock properties.[6,11-14] In particular, convolutional neural networks (CNNs) have shown improved recovery of high-frequency components and robustness to moderate noise relative to purely descent-based or strongly linearized inversions.[7-11] At the same time, physics-guided DL—via architectural design or loss regularization, where additional physics-based penalty terms are introduced into the loss function to enforce consistency with rock-physics or wave-propagation principles—has emerged to couple data-driven representations with geophysical constraints, improving interpretability and generalization.[15-19] Physics-guided DL incorporates physical priors into data-driven models through architectural design or loss regularization, whereas physics-constrained inversion directly embeds forward modeling into the inversion. Generative and uncertainty-aware models further help capture non-uniqueness and quantify reliability.[20] Nevertheless, many DL applications still rely on post-stack inputs or treat pre-stack traces independently, underutilizing the angle-dependent reflectivity and thin-bed interference information critical to coal reservoirs. Moreover, most existing physics-guided CNNs or generative adversarial network (GAN)-based inversions are developed for generalized impedance prediction or post-stack attributes, with limited adaptability to thin coal seams where reflections overlap within a single lobe, and signal-to-noise ratio (SNR) is intrinsically low. Similarly, attention-fusion strategies enhance multi-scale features but often lack explicit mechanisms to enforce rock-physics plausibility, which constrains their reliability in complex CBM settings.

To bridge these gaps, we develop a two-branch CNN (TB-CNN) for direct ν prediction in coal-bearing gas reservoirs. The first branch ingests elastic impedance (EI)-inverted Vp, Vs, and ρ to provide stable, physics-consistent large-scale trends that anchor the solution space, while the second branch processes small-, medium-, and large-angle partial stacks that preserve high-resolution waveform variations sensitive to thin-bed effects. The two streams are fused in a shared latent space with soft rock-physics regularization derived from well-log relationships, enabling a nonlinear mapping directly to ν and mitigating error propagation from intermediate inversions. Conceptually, this design differs from prior physics-guided CNNs or GAN-based inversions in three respects: (i) it explicitly integrates angle-dependent reflectivity information often ignored in post-stack or single-trace DL workflows; (ii) it embeds empirical rock-physics constraints as soft priors, improving interpretability and avoiding geologically implausible predictions common in unconstrained GAN outputs; and (iii) it is tailored to thin-bed coal seams where conventional inversions fail due to wavelet interference. These design choices allow TB-CNN to leverage the complementary strengths of physics-derived attributes and multi-angle data-driven features: the former stabilizes training by providing low-frequency, physically consistent background trends, while high-resolution information is primarily recovered from multi-angle seismic data through nonlinear feature learning, thereby enhancing vertical resolution and lateral detectability beyond what linearized inversions typically deliver.[7,11,21]

We validate the framework on field data from the northern Yishan Slope of the Ordos Basin, where the 8# coal seam (average thickness ~8.3 m) exhibits a strong impedance contrast but a thin-bed seismic response embedded within a single lobe. Using three-dimensional (3D) pre-stack seismic data and a suite of wells for supervision and blind testing, we show that TB-CNN-predicted ν volumes delineate seam boundaries more sharply, better capture gradual thickness variations and localized thinning, and correlate more closely with log-

derived ν than conventional EI-based three-parameter inversions followed by ratio computation. Horizon-slice analysis further indicates superior lateral continuity and sensitivity to subtle heterogeneity. These results substantiate that integrating physics-informed trends with multi-angle data-driven features can reduce dependence on linearized approximations and smooth initial models, thereby enhancing the reliability of ν prediction for CBM reservoir characterization in low SNR, thin-bed environments.[1,2]

## 2. Methods

This study focuses on predicting elastic parameters in coal-bearing gas reservoirs, which are often characterized by thin beds, strong impedance contrasts, and low SNR. These geological and seismic conditions make inversion particularly challenging. Conventional pre-stack inversion methods rely on linearized forms of the Zoeppritz equations, where high-frequency information is frequently suppressed due to smoothing constraints and approximation errors. In practice, the smooth initial model, together with the regularization applied during inversion, further attenuates thin-bed responses and weak reflections. As a result, conventional approaches typically struggle to preserve fine-scale heterogeneity and fail to accurately delineate thin coal seams. This limitation is especially evident when predicting ν, which is highly sensitive to subtle variations in both Vp and Vs.

To mitigate these limitations, we propose the TB-CNN model that integrates physics-constrained inversion outputs with high-resolution seismic data in a unified learning framework, shown in Figure 1. The first branch (Branch 1) ingests the Vp, Vs, and ρ volumes obtained through EI inversion, providing stable, large-scale trends consistent with rock-physics principles. The second branch (Branch 2) directly utilizes stacked pre-stack seismic data from three incidence-angle ranges (small, medium, and large), thereby retaining high-frequency waveform details and thin-bed sensitivity without dependence on inversion linearization.

Feature representations extracted by the two branches are fused within a shared latent space, where empirical rock-physics constraints derived from well-log analysis are implicitly embedded to ensure geological plausibility. The fused features are subsequently mapped to the target elastic parameters through a sequence of fully connected (FC) layers. By jointly leveraging the stability and interpretability of physics-based trends from Branch 1 and the fine-scale, noise-resilient information from Branch 2, the proposed TB-CNN reduces reliance on linearized forward models, mitigates bias introduced by smooth initial models, and enhances the detectability of subtle elastic variations in complex coal-bearing gas reservoirs.

### 2.1. Branch 1 design: Physics-constrained elastic parameter input

The first branch of the proposed TB-CNN incorporates physics-based prior information obtained from conventional EI inversion, using the inverted Vp, Vs,

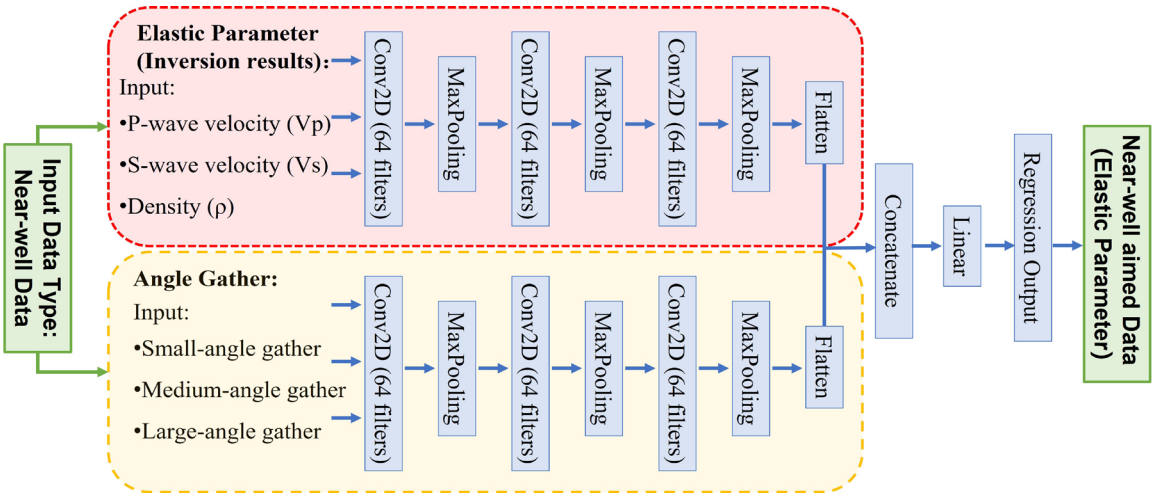



**Figure 1.** Schematic workflow of the proposed two-branch convolutional neural network-based inversion methodology
Abbreviation: Conv2D: Two-dimensional convolution.

and ρ volumes as inputs (Figure 2). These parameters are estimated through a multi-angle inversion procedure grounded in the linearized Zoeppritz equations and the EI formulation, thereby ensuring consistency with established rock-physics relationships.

To extend acoustic impedance to non-normal incidence, the concept of EI was introduced, which expresses the angle-dependent reflectivity R(θ) in an impedance-like form consistent with the Aki–Richards approximation.[22] Given pre-stack seismic reflectivity R(θ) at incidence angle θ, the linearized Aki–Richards equation can be written as **Equation 1**:

$$R(\theta) \approx A + B\sin^2\theta + C(\tan^2\theta - \sin^2\theta) \tag{1}$$

where A, B, and C depend on the fractional contrasts in Vp, Vs, and ρ across an interface, as shown in **Equations 2–4**:

$$A = \frac{1}{2}\left(\frac{\Delta V_P}{V_P} + \frac{\Delta\rho}{\rho}\right) \tag{2}$$

$$B = \frac{\Delta V_p}{V_p} - 2\left(\frac{V_s}{V_p}\right)^2\left(\frac{\Delta V_s}{V_s} + \frac{\Delta\rho}{\rho}\right) \tag{3}$$

$$C = \frac{1}{2}\frac{\Delta V_p}{V_p} \tag{4}$$

For a reference angle $\theta_0$, the EI ($\theta_0$) is defined as **Equation 5**:

$$EI(\theta_0) = V_p^{a(\theta_0)} V_s^{b(\theta_0)} \rho^{c(\theta_0)} \tag{5}$$

where K = $(Vs/Vp)^2$ is the shear-to-compressional velocity ratio, and by matching the logarithmic contrast of EI ($\theta_0$) with the Aki–Richards expression, the exponents are obtained as **Equations 6–8**:

$$a(\theta_0) = 1 + \tan^2\theta_0 \tag{6}$$

$$b(\theta_0) = -8K\sin^2\theta_0 \tag{7}$$

$$c(\theta_0) = 1 - 4K\sin^2\theta_0 \tag{8}$$

Performing EI inversion over multiple reference angles enables the joint recovery of Vp, Vs, and ρ volumes that preserve large-scale geological trends and conform to rock-physics constraints derived from well logs. However, such inversions remain susceptible to smoothing effects and linearization errors, particularly in thin coal seams and highly heterogeneous zones. The EI formulation thus preserves the physical consistency between angle-dependent reflectivity and elastic parameters, providing a stable and interpretable input for the physics-constrained branch of the TB-CNN.

In the TB-CNN framework, these three inverted volumes are concatenated into a three-channel input tensor, as shown in **Equation 9**:

$$X_1 \in \mathrm{R}^{H\times W\times 3} = [V_p, V_s, \rho] \tag{9}$$

where *H* and *W* denote the spatial dimensions in inline–crossline or time–depth coordinates. Branch 1 applies

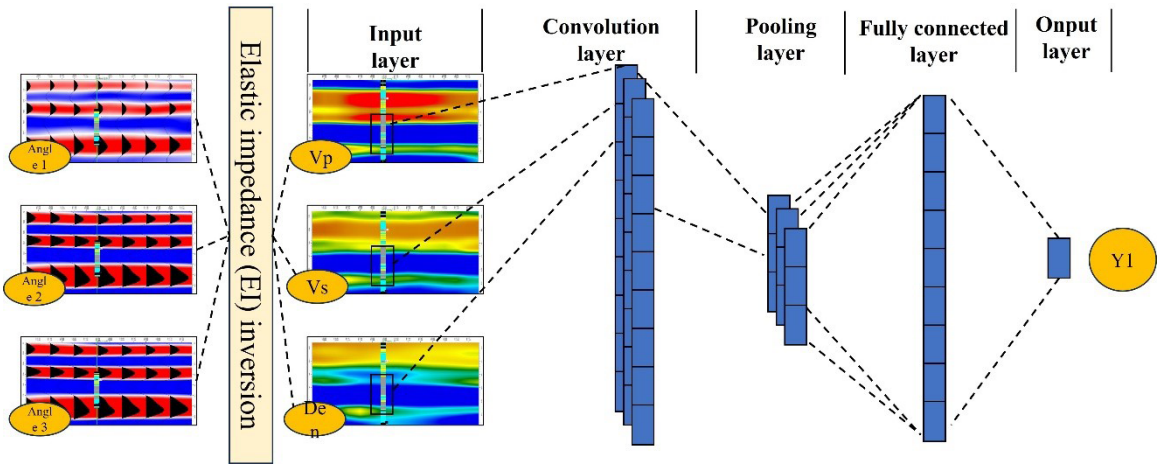



**Figure 2.** Branch 1 convolutional structure diagram
Abbreviations: ρ: Density; Vp: P-wave velocity; Vs: S-wave velocity.

a sequence of two-dimensional convolutional layers, as shown in **Equation 10**:

$$F_1^{(l)} = \sigma\left(W^{(l)} * F_1^{(l-1)} + b^{(l)}\right), l = 1, \ldots, L_1 \quad (10)$$

where $F_1^{(l)}$ = X1, W (1) and b (1) are the trainable convolution kernels and biases, and σ(·) is a nonlinear activation function (e.g., Rectified Linear Unit [ReLU]). These operations progressively extract multi-scale spatial features that encode the physics-consistent elastic property patterns from the inversion-derived parameters.

After the convolutional feature extraction, the final feature maps of Branch 1 are flattened and projected into a latent representation, as shown in **Equation 11**:

$$z_1 = \text{Flatten}\left(F_1^{(L_1)}\right) W_z + b_z \quad (11)$$

where $W_z$ and $b_z$ are trainable parameters. This latent vector $z_1$ encapsulates the physics-consistent information from inversion-derived Vp, Vs, and ρ, and serves as one of the two complementary inputs to the feature-fusion module.

By leveraging EI-derived Vp, Vs, and ρ as inputs, Branch 1 preserves physically meaningful structural trends while allowing subsequent network layers to learn data-driven corrections for systematic inversion biases. The resulting feature embeddings are later fused with high-resolution, waveform-sensitive features from Branch 2 to improve thin-bed detectability and robustness under low-SNR conditions.

### 2.2. Branch 2 design: Multi-angle seismic input (data-driven)

The second branch of the proposed TB-CNN directly ingests three incidence-angle seismic volumes and learns a purely data-driven mapping from raw seismic amplitudes to the target elastic parameter(s), without invoking any explicit AVO formulation. The three input volumes are obtained from partial-angle stacks (small, medium, and large incidence angles) extracted after standard pre-stack processing, including velocity analysis, NMO correction, and angle-domain binning. Each angle stack captures complementary reflectivity patterns arising from different incidence angles, enabling the network to exploit both near-ofset continuity and far-ofset contrast sensitivity. These three stacks are spatially co-registered with the inversion volumes in Branch 1, ensuring consistent supervision windows (Figure 3).

The three partial-angle stacks are concatenated to form a three-channel input tensor, as shown in **Equation 12**:

$$X_2 \in \mathrm{R}^{H \times W \times 3} = [S_g, S_m, S_l] \quad (12)$$

where $S_g$, $S_m$, and $S_l$ denote the small-, medium-, and large-angle stacks, respectively, and $H \times W$ corresponds to the spatial sampling in inline/crossline or trace/time dimensions.

Branch 2 applies a series of two-dimensional convolutional layers designed to extract short-lag waveform features and thin-bed interference signatures. Let $F_2^{(0)} = X_2$ be the input to the branch. The convolutional

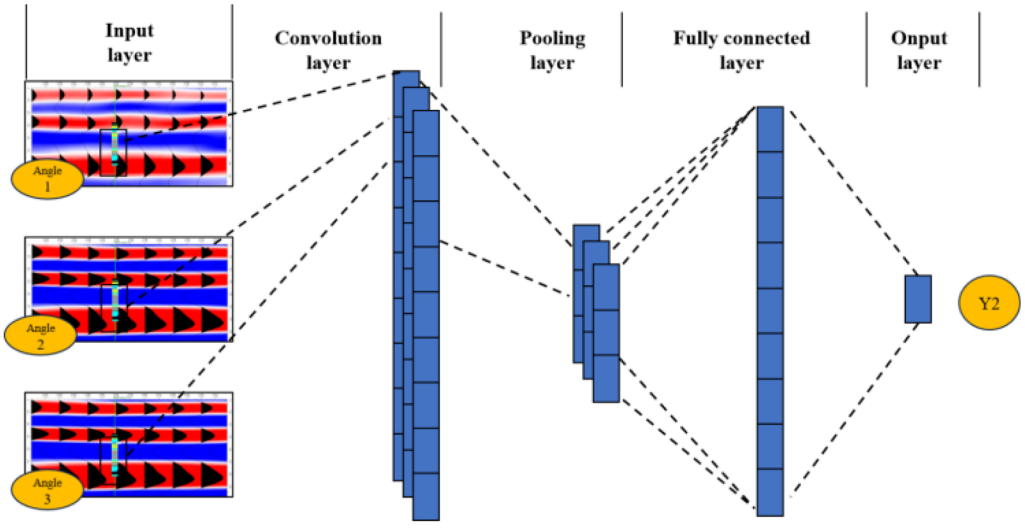



**Figure 3.** Branch 2 convolutional structure diagram

feature extraction proceeds as **Equation 13**:

$$F_2^{(l)} = \sigma\left(W^{(l)} * F_2^{(l-1)} + b^{(l)}\right), l = 1, \ldots, L_2 \quad (13)$$

where $W^{(l)}$ and $b^{(l)}$ are the learnable kernels and biases, * denotes the two-dimensional convolution operator, and σ(·) is a nonlinear activation (e.g., ReLU). Kernel sizes are chosen to emphasize short-period oscillations along the temporal axis (e.g., 1 × 2 or 3 × 3), which are particularly sensitive to thin-bed tuning effects. Pooling layers may be interleaved to reduce temporal redundancy and improve robustness to random noise.

After the final convolutional block, the feature maps are flattened and linearly projected into a latent representation, as shown in **Equation 14**:

$$z_2 = \text{Flatten}(F_2^{(L_2)})W_z + b_z \quad (14)$$

where $W_z$ and $b_z$ are trainable parameters. This vector $z_2$ encapsulates high-resolution reflectivity information derived directly from the multi-angle seismic data and complements the physics-consistent $z_1$ produced by Branch 1. Both $z_1$ and $z_2$ are subsequently fused in the shared latent space to generate the final elastic parameter predictions.

## 2.3. Feature fusion and output prediction

After the independent processing of the physics-informed branch (Branch 1) and the purely data-driven seismic branch (Branch 2), their respective embeddings $z_1$ and $z_2$ are concatenated to form a joint latent representation, as shown in **Equation 15**:

$$z = \text{Concat}[z_1, z_2] \in \mathrm{R}^{d_1+d_2} \quad (15)$$

where $d_1$ and $d_2$ denote the dimensionalities of $z_1$ and $z_2$, respectively. This concatenation preserves complementary information: $z_1$ encodes large-scale geological trends and physically interpretable elastic responses from inversion results, whereas $z_2$ provides fine-scale seismic patterns and thin-bed interference signatures extracted directly from waveform data.

### *2.3.1. Joint feature refinement*

The fused latent vector z is subsequently passed through a sequence of FC layers to enable nonlinear cross-branch interaction and hierarchical abstraction. Each FC layer is followed by batch normalization (BN) and a nonlinear activation function (ReLU) to improve convergence stability and prevent gradient saturation. Dropout layers (drop probability 0.2–0.3) are optionally applied after BN to enhance generalization. The transformation is expressed as **Equation 16**:

$$h^{(l)} = \sigma\left(\text{BN}\left(w^{(l)}h^{(l-1)} + b^{(l)}\right)\right), l = 1, \ldots, L \quad (16)$$

with $h^{(0)} = z$, $W^{(l)}$ and $b^{(l)}$ are trainable parameters, and σ(·) denotes the activation.

### *2.3.2. Prediction layer*

The final FC layer maps the refined feature vector $h$ ($L$) to the predicted elastic parameter(s) $\hat{y}$, as shown in **Equation 17**:

$$\hat{y} = w^{(o)}h^{(L)} + b^{(o)} \quad (17)$$

where $\hat{y}$ can be a scalar (e.g., a single elastic modulus) or a vector (e.g., Vp, Vs, and ρ jointly). Model training minimizes the mean squared error (MSE) loss between network predictions and well-log-derived ground truth, as shown in **Equation 18**:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\hat{y}^{(i)} - y^{(i)}\right\|_2^2 \quad (18)$$

where $N$ is the number of training samples.

### *2.3.3. Advantages of dual-branch fusion*

The proposed fusion architecture leverages the complementary nature of the two branches:

(i) Physics-informed branch (Branch 1): Provides large-scale structural context and physically interpretable elastic attributes, ensuring that predictions remain geophysically consistent.

(ii) Seismic-driven branch (Branch 2): Supplies high-resolution, noise-resilient features that capture subtle waveform and stratigraphic variations not recoverable by conventional inversion.

By unifying these representations within a shared latent space, the network achieves enhanced robustness to noise, improved thin-bed resolution, and increased predictive accuracy under low SNR conditions.

## 2.4. Network architecture details

To ensure reproducibility, the complete TB-CNN architecture is summarized in Table 1. Both branches adopted a symmetric design with two convolutional blocks, each followed by batch normalization and ReLU activation. Pooling layers were interleaved to enhance generalization, while the fusion stage employed two FC layers prior to the final regression output.

Although the TB-CNN employs two parallel branches, each branch contains only two convolutional blocks with moderate kernel dimensions and channel counts, forming a lightweight architecture with fewer than 1.2 million trainable parameters. In addition, the network operates on fixed-size sliding windows centered on the target coal-bearing interval rather than the entire 3D volume. As only three representative partial-angle stacks (small, medium, and large) were used, the effective data volume is about three times that of a post-stack dataset but far smaller than the full pre-stack gathers. Together, these design choices ensure that the computational cost remains moderate and the network can be efficiently trained and deployed on standard graphics processing unit hardware.

### 2.5. Soft rock-physics regularization

To ensure geological plausibility, we embedded empirical rock-physics constraints into the learning process as a soft regularization term. Specifically, we constructed a penalty $L_{phys}$ based on the deviation of predicted elastic relationships from well-log-derived regressions, such as **Equation 19**:

$$L_{phys} = \frac{1}{M}\sum_{j=1}^{M}\left(\alpha\left(\frac{V_s}{V_p}\right)_j - \beta\rho_j - \gamma\right)^2 \quad (19)$$

where $(Vs/Vp)_j$ and $\rho_j$ denote the predicted parameters at location j, while α, β, and γ are the regression coefficients fitted from the well-log data. This formulation penalizes predictions that violate empirical Vp–Vs–ρ relationships, thereby reducing the likelihood of nonphysical solutions.

The final loss function is a weighted sum of data misfit and physics-based regularization, as shown in **Equation 20**:

$$L = L_{MSE} + \lambda L_{phys} \quad (20)$$

where $L_{MSE}$ is the conventional mean squared error, $L_{phys}$ is the soft rock-physics penalty, and λ is a hyperparameter (0.1–0.5 in our experiments) that balances accuracy and physical consistency. It should be noted that the proposed regularization represents an empirical rock-physics consistency constraint rather than a hard physical constraint derived from forward modeling or wave-equation operators.

This strategy differs from hard constraints in that deviations from empirical trends are not strictly forbidden, but gradually penalized. As a result, the network retains flexibility to adapt to local heterogeneities while preserving global consistency with rock-physics knowledge.

## 3. Field data test

### 3.1. Study area and data

The study area is located in the northern segment of the Yishan Slope, Ordos Basin, spanning Wushen Banner in Inner Mongolia and Yuyang District and Mizhi County in Shaanxi Province (Figure 4). The structural setting is characterized by a relatively stable slope belt, where Carboniferous–Permian coal-bearing strata are widely developed. The target reservoir was the 8# coal seam of the Shanxi Formation, with a thickness ranging from 5.8 m to 9.7 m (average 8.3 m). This seam serves as the primary production horizon for CBM development in the region and is distinguished by a pronounced acoustic impedance

**Table 1. Detailed architecture of the proposed TB-CNN**

| Layer | Branch | Input dimension | Kernel/Stride | Output dimension |
|---|---|---|---|---|
| Conv1 | Branch 1 (EI inputs) | H × W × 3 | 3 × 3/1 | H × W × 32 |
| Conv2 | Branch 1 | H × W × 32 | 3 × 3/1 | H × W × 64 |
| Pooling | Branch 1 | H × W × 64 | 2 × 2/2 | $\frac{H}{2}\times\frac{W}{2}\times 64$ |
| Conv1 | Branch 2 (Angle stacks) | H × W × 3 | 3 × 3/1 | H × W × 32 |
| Conv2 | Branch 2 | H × W × 32 | 3 × 3/1 | H × W × 64 |
| Pooling | Branch 2 | H × W × 64 | 2 × 2/2 | $\frac{H}{2}\times\frac{W}{2}\times 64$ |
| Flatten | Both branches | - | - | $1 \times (d_1 + d_2)$ |
| FC1 | Fusion | $d_1 + d_2$ | - | 128 |
| FC2 | Fusion | 128 | | 64 |
| Output | Fusion | 64 | | 1 (Poisson's ratio) |

Note: H × W denotes the input patch dimension, which is consistent across both branches.
Abbreviations: Conv: Convolution; FC: Fully connected; TB-CNN: Two-branch convolutional neural network.

contrast with surrounding lithologies.

To illustrate the seismic response and geological continuity of the target seams, well–seismic ties from three representative boreholes—Shuang50–45, Shuang133, and Shuang48–45—are shown in Figure 5. In all three wells, the 8# coal seam exhibited a strong impedance contrast with the surrounding mudstone–sandstone–limestone assemblage, producing a high-amplitude composite reflection with polarity reversal indicative of thin-bed tuning. This consistent response across multiple wells validates the seam interpretation and provides key calibration data for pre-stack inversion and TB-CNN training.

The regional continuity of the coal-bearing sequence is further demonstrated by the cross-well seismic section in Figure 6. Four key composite reflection events were identified. Given that the 5# and 8# coal seams are sub-wavelength thin beds (typically thickness < λ/8), their seismic responses are dominated by tuning interference rather than by isolated interface reflections. Consequently, the observed positive and negative phases represent composite waveforms generated by constructive and destructive superposition of closely spaced boundaries, and should not be interpreted as discrete reflections from the seam roof and floor. In this study, these composite events were used only to delineate a consistent inversion window for subsequent elastic parameter prediction, rather than to infer detailed coal–seam interface geometries.

Within this interval, the 8# seam maintains remarkable lateral continuity among the three wells, though thin-bed interference blurs the roof and floor reflections, limiting vertical resolution. These features highlight both the geological importance of the 8# seam as a major CBM reservoir and the inherent limitations of conventional pre-stack inversion in resolving thin coal seams.

To address these challenges, this study applied the proposed TB-CNN to the field dataset. By integrating physics-constrained inversion products with multi-angle seismic inputs, the TB-CNN was designed to enhance thin-bed resolution and improve ν inversion for CBM reservoir characterization.

### 3.2. Network training and direct Poisson's ratio prediction

Poisson's ratio is a highly diagnostic elastic parameter for coal-bearing formations, as it reflects both lithological variations and fluid content. In rock physics, ν can be expressed in terms of the Vs/Vp ratio as shown in **Equation 21**:

$$\nu = \frac{V_p^2 - 2V_s^2}{2(V_p^2 - V_s^2)} = \frac{1}{2} \cdot \frac{1 - 2\left(\frac{V_s}{V_p}\right)^2}{1 - \left(\frac{V_s}{V_p}\right)^2} \tag{21}$$

which provides the physical basis for deriving ν from well-log measurements. In this study, however, ν was not computed from post-inverted seismic velocities. Instead, it was directly predicted by the TB-CNN, with the above relationship serving only to define the supervision labels and guide physical interpretation.

In the context of CBM exploration, accurate estimation of ν is particularly important because it governs rock brittleness, fracture propagation behavior, and gas-bearing potential. Thin coal seams present a persistent challenge for conventional EI inversion: their seismic responses are commonly embedded within a single tuning lobe, and the resulting Vp and Vs estimates are often over-smoothed. Consequently, ν computed from EI-derived velocities suffers from blurred seam boundaries and limited sensitivity to thickness variations and local thinning. To address these limitations, we adopted a direct-inversion strategy in which TB-CNN is trained to predict ν directly from multi-source seismic inputs, thereby avoiding error accumulation associated with sequential velocity estimation.

The TB-CNN integrates two complementary input domains: (i) EI-inverted elastic parameters (Vp, Vs, and ρ) as physics-consistent prior information (Branch 1), and (ii) partial-angle seismic stacks as high-resolution reflectivity inputs (Branch 2). Figure 7A–C shows the small-, medium-, and large-angle stacks used as Branch 2 inputs, while Figure 8A–C displays the EI-inversion products serving as Branch 1 inputs. This dual-input design enables the network to preserve large-scale elastic trends while exploiting high-frequency waveform variations associated with thin-bed tuning.

The training dataset was constructed using two wells with reliable log coverage across the 8# coal seam, located at approximately 1,400 ms two-way travel time. Around each well, fixed-size two-dimensional patches were extracted in both inline and crossline directions, with patch size and stride selected to balance sample diversity and spatial continuity. Each patch was centered on the well trajectory to ensure alignment between seismic inputs and the supervision target. The Branch 1 input $X1 \in RH \times W \times 3$ contains Vp, Vs, and ρ, while the Branch 2 input $X2 \in RH \times W \times 3$ consists of the three angle stacks.

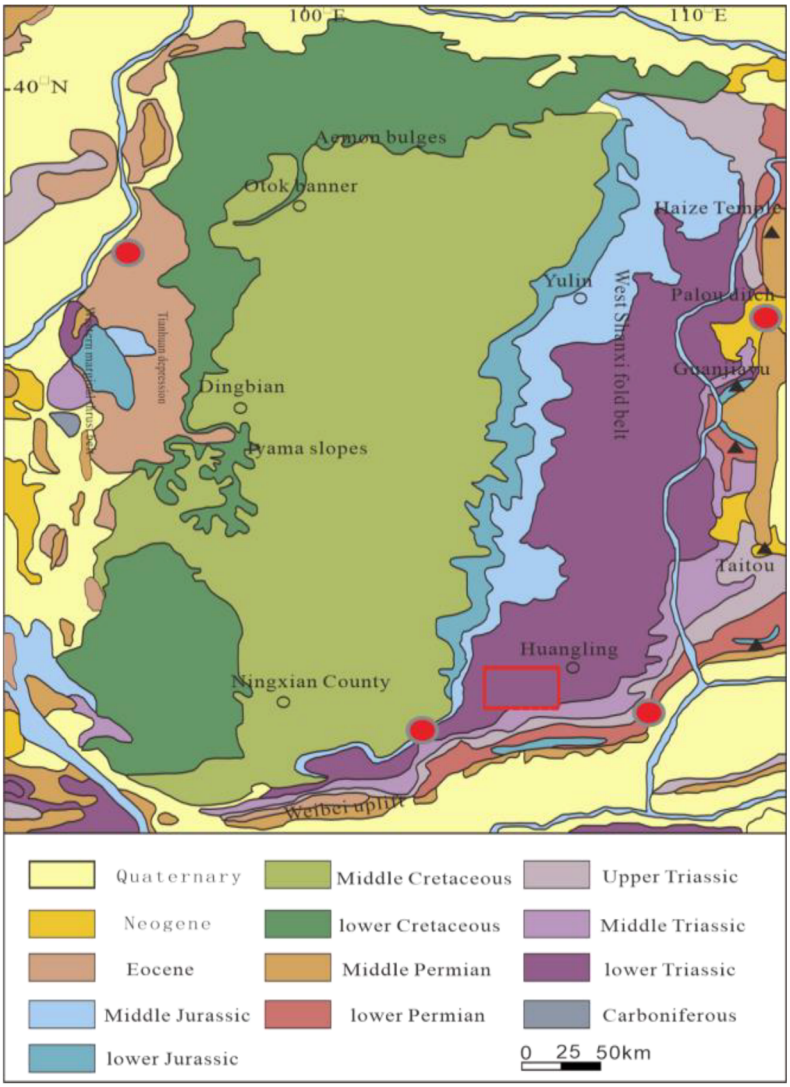



**Figure 4.** Location map of the study area in the northern Yishan Slope of the Ordos Basin. The red rectangle marks the three-dimensional seismic survey boundary, and red circles indicate the distribution of wells used for inversion and two-branch convolutional neural network training.

The supervision label for each patch is the log-derived ν, as shown in **Equation 22**,

$$\nu^{\log} = \frac{1}{2}\left(1 - 2\left(\frac{\nu^{\log}}{\nu_p^{\log}}\right)^2\right) \tag{22}$$

which is depth–time converted and wavelet-aligned to the seismic domain. All input channels are standardized using the mean and variance of the training set to ensure numerical stability.

The TB-CNN was trained to directly regress ν via a fused latent representation, as shown in **Equation 23**:

$$\hat{\nu} = f_\theta(X_1, X_2) \tag{23}$$

with the network output constrained to a geologically realistic interval of ν ∈ (0.30, 0.38) based on well-log statistics and regional lithology. This constraint was applied during training to stabilize regression rather than enforcing uniform or constant values in the prediction stage.

The objective function combines the MSE loss with a soft physics-guided regularization term, as shown in **Equation 24**,

$$L = L_{\text{MSE}} + \lambda L_{\text{phys}} \tag{24}$$

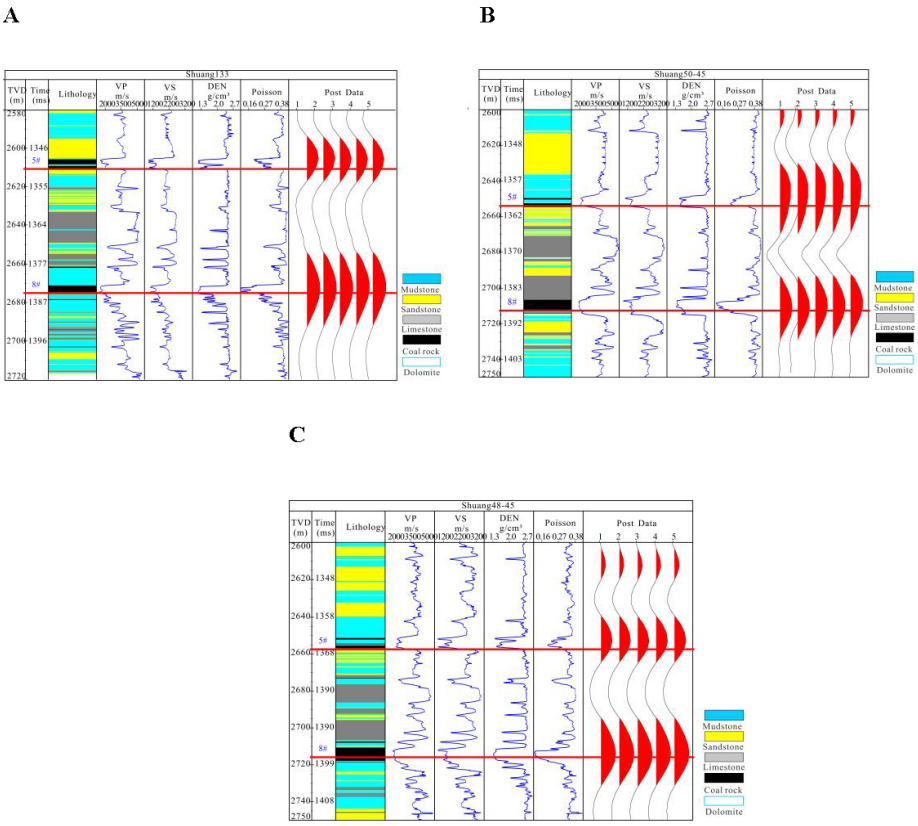



**Figure 5.** Well-to-seismic ties of three representative boreholes in the study area. (A) Shuang50–45, (B) Shuang133, and (C) Shuang48–45. The 8# coal seam, delineated by the red horizontal line, shows a pronounced impedance contrast with the adjacent mudstone, sandstone, and limestone layers, producing a strong composite reflection with polarity reversal indicative of thin-bed tuning. The consistent seismic signature across the three wells confirms the geological reliability of the 8# seam interpretation and provides essential calibration for pre-stack inversion and two-branch convolutional neural network training. The positions of both the 5# and 8# coal seams are consistently indicated across all figures to ensure unified stratigraphic reference.

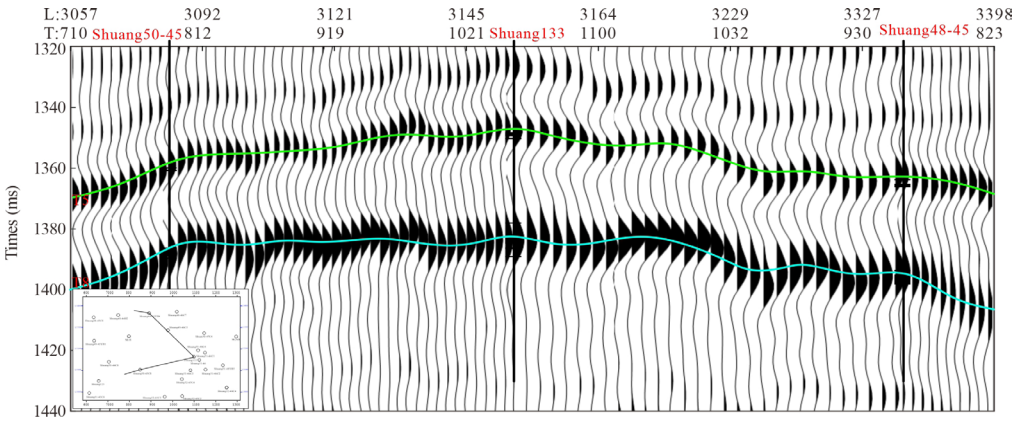



**Figure 6.** Cross-well seismic section illustrating the lateral continuity of the coal-bearing sequence. Four key composite reflection events were identified: T5 and T8 correspond to the dominant tuning-controlled composite responses associated with the 5# and 8# coal seams. These reflections do not represent discrete roof or floor interfaces, and their positive/negative phases should not be interpreted as top/bottom seam reflections. Instead, they arise from constructive and destructive waveform superposition caused by thin-bed tuning (thickness < λ/8 at the dominant frequency). The figure demonstrates the regional continuity of tuning responses and defines a physically reasonable inversion window for subsequent elastic parameter prediction.

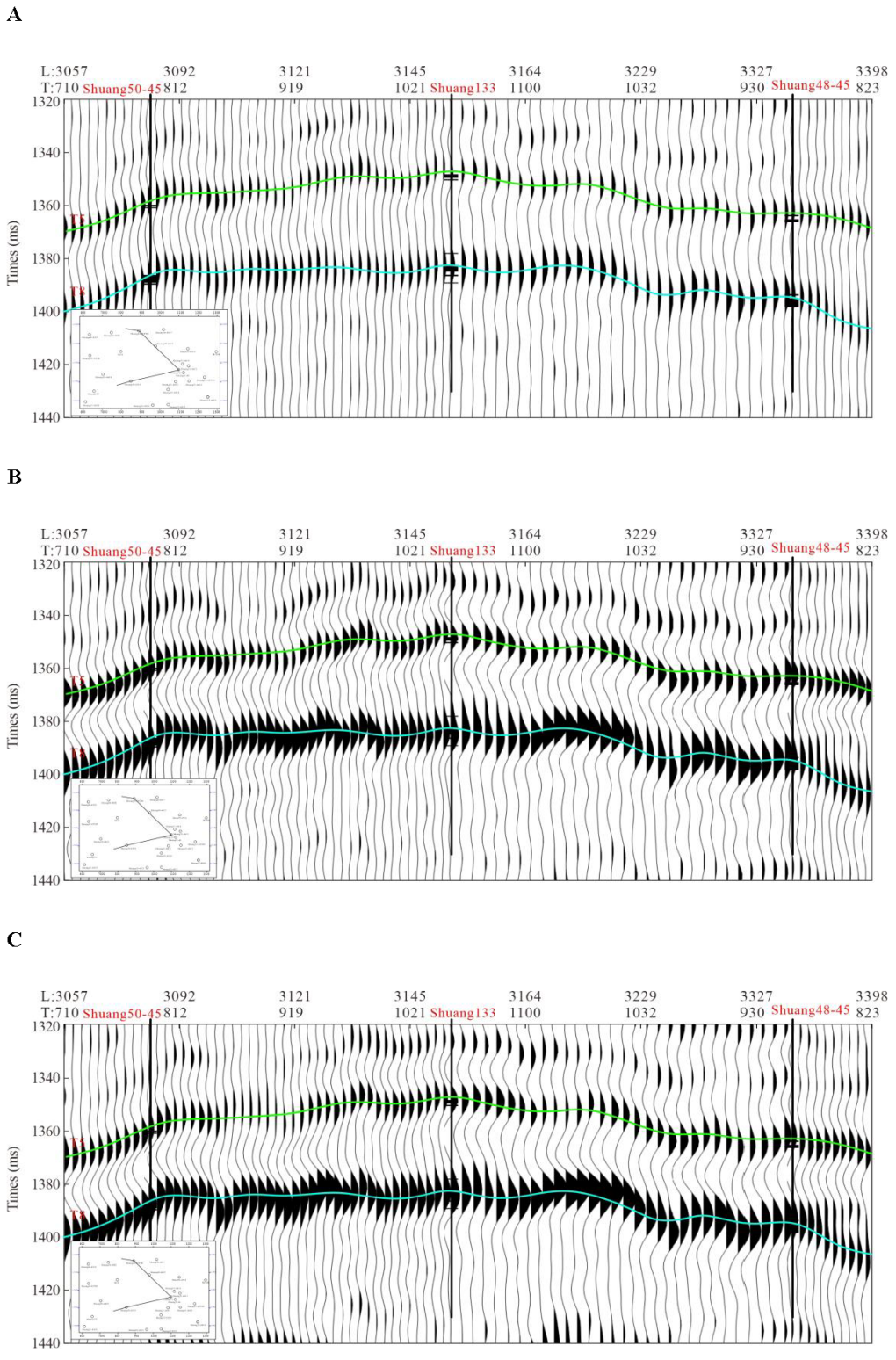



**Figure 7.** Partial-angle stacked seismic profiles used as Branch 2 inputs. (A) Small-angle stack, (B) medium-angle stack, and (C) large-angle stack.

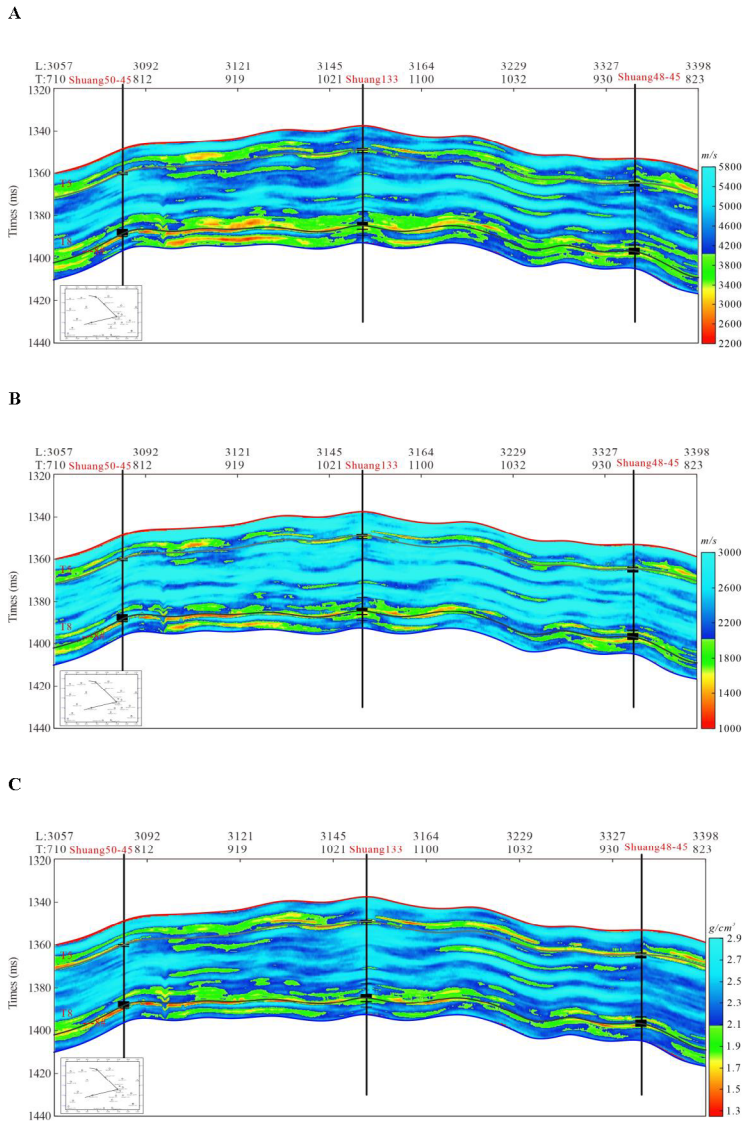



**Figure 8.** Elastic impedance-inversion products used as Branch 1 inputs. (A) P-wave velocity (Vp), (B) S-wave velocity (Vs), and (C) density (ρ).

which enforces consistency between the predicted ν and empirical rock-physics relationships implied by the input Vp, Vs, and ρ volumes. Rather than constraining velocities explicitly, this regularization penalizes ν predictions that violate lithology-consistent elastic trends inferred from well logs, thereby improving stability and interpretability.

Optimization was performed using the Adam algorithm with an initial learning rate of $1 \times 10^{-4}$, a batch size of 32, and a weight decay of $1 \times 10^{-5}$. Batch normalization and dropout were applied to mitigate overfitting, and early stopping was employed based on validation loss. Data splitting was performed strictly at the well level to avoid spatial leakage, so that blind wells provided an independent assessment of generalization. Convergence was typically achieved within 80–100 epochs, with stable training and validation loss curves indicating robust learning (Figure 9).

Once trained, the TB-CNN was applied to the full 3D seismic volume in a sliding-window manner to generate a continuous ν cube for reservoir characterization. An example inline ν section predicted by TB-CNN is shown in Figure 10, which illustrates the direct network output used in subsequent analyses.

The TB-CNN result (Figure 10) reveals laterally continuous composite reflection bands associated with the 5# and 8# coal seams. These bands represent

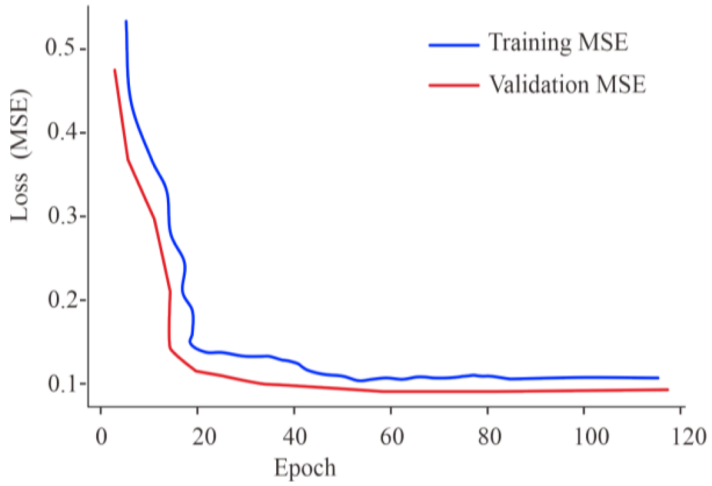



**Figure 9.** Training and validation loss curves of the two-branch convolutional neural network, showing stable convergence within approximately 100 epochs
Abbreviation: MSE: Mean squared error.

tuning-dominated seismic responses generated by thin coal layers, rather than isolated reflections from individual roof or floor interfaces. Despite the presence of tuning interference, the TB-CNN preserves the relative stratigraphic position and lateral continuity of the coal-bearing intervals, allowing thickness variations and local thinning trends to be recognized. Within the coal seams, ν values were generally elevated in this study area, from approximately 0.30 to 0.38, reflecting the reduced shear rigidity and higher compressibility of coal. Between the 5# and 8# coal seams, ν values showed partial overlap with those of the surrounding sandstone–limestone interbeds, which is expected in interbedded sedimentary sequences. The spatial distribution and relative contrasts of ν were consistent with the log-derived elastic properties, indicating that the TB-CNN prediction captured the first-order lithological variations within the coal-bearing sequence. Based on well-log data from three representative wells in the study area, the ν of the 8# coal seam typically ranged from 0.30 to 0.38. This range aligns with typical

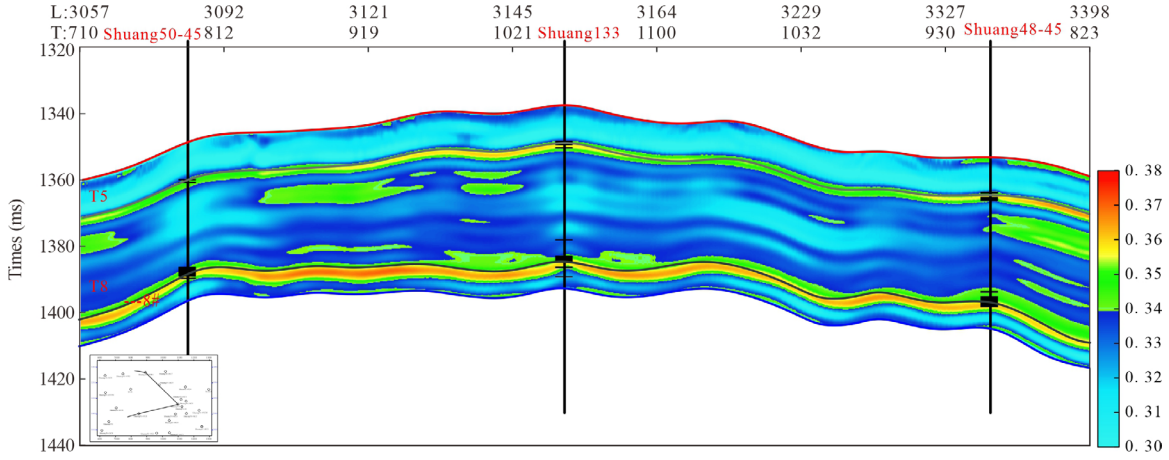



**Figure 10.** Inline profile of Poisson's ratio predicted by the proposed two-branch convolutional neural network. Coal-bearing intervals are characterized by relatively elevated Poisson's ratio values, forming laterally continuous composite bands controlled by thin-bed tuning effects. These responses do not represent isolated roof or floor reflections, but rather the effective elastic behavior of thin coal seams. In contrast, comparatively lower or transitional Poisson's ratio values were observed in the interbedded sandstone–limestone units between coal seams.

values reported for coal seams under similar geological conditions, characterized by low shear rigidity and higher compressibility.

### 3.3. Results comparisons

To further demonstrate the effectiveness of the proposed TB-CNN, inline sections of the inverted ν were compared with results obtained from two reference methods: (i) conventional EI inversion followed by Vp–Vs–based ν computation (Figure 11), and (ii) a purely data-driven single-branch CNN (SB-CNN) that directly predicts ν from seismic data (Figure 12). The corresponding TB-CNN result is shown in Figure 10.

The side-by-side comparison highlights pronounced differences in the ability of the three approaches to characterize thin coal seams under strong impedance contrast and tuning-dominated conditions. Relative to the EI- and SB-CNN-based results, the TB-CNN section (Figure 10) provided a stable and geologically coherent reference for seam continuity and thickness variation.

The EI-derived Poisson's ratio section (Figure 11) was dominated by over-smoothing effects, which blurred coal–seam boundaries and suppressed thin-bed signatures as a result of linearized inversion assumptions and band-limited seismic input. In contrast, the SB-CNN result (Figure 12) partially enhanced thin-bed visibility but suffered from local amplitude fluctuations, discontinuous patches, and reduced lateral continuity, reflecting the inherent instability of purely data-driven inversion when physics-based constraints are absent.

To further validate the inversion reliability, near-well comparisons were performed at three representative wells (Shuang50–45, Shuang133, and Shuang48–45), where each profile included a lithology column, log-derived ν, and inversion results from EI inversion, SB-CNN, and the proposed TB-CNN (Figure 13).

The TB-CNN predictions (red curves) showed the highest agreement with the log-derived ν. They accurately delineated the top and bottom of the coal seams, preserved thin-bed variations, and reproduced both the magnitude and trend of ν across lithological transitions. In contrast, EI inversion yielded overly smooth curves with weak sensitivity to interbedded coal layers and limited correspondence to lithological boundaries. The SB-CNN improved local resolution but still exhibited unstable amplitudes and diminished contrast, resulting in poorer alignment with both the log-derived ν and the lithology column.

A comparative analysis of the horizon slices (Figures 14–16) extracted along the 8# coal seam further highlights the advantages of the proposed TB-CNN in characterizing lateral heterogeneity and continuity of coal-bearing reservoirs. The TB-CNN prediction (Figure 14) exhibited a laterally coherent ν distribution with clearly resolved localized anomalies, effectively capturing thin-seam variability and subtle lateral changes that are consistent with the known geological framework and gas-producing well distribution. In contrast, the horizon slice derived from conventional EI inversion (Figure 15) appeared overly smoothed, with weakened lateral contrasts and blurred

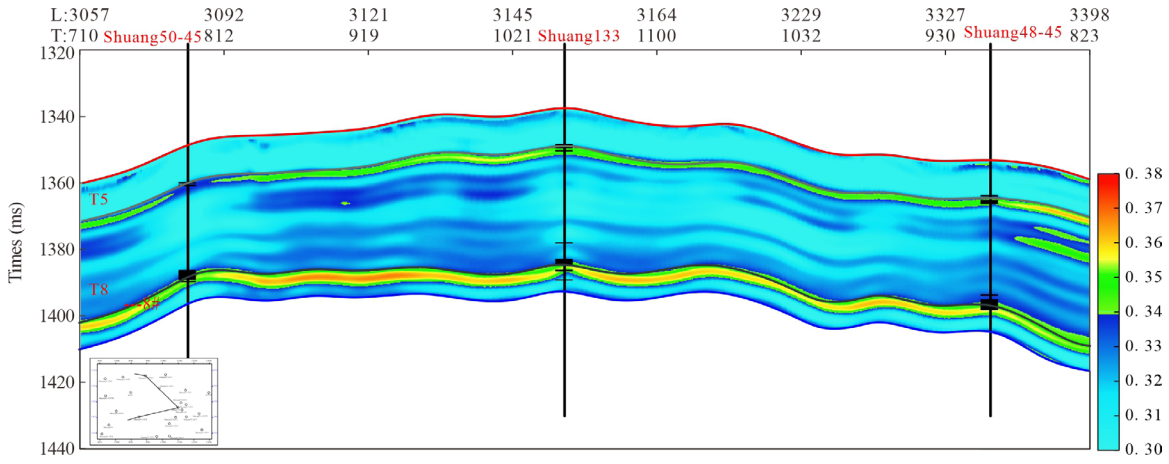



**Figure 11.** Inline profile of Poisson's ratio derived from conventional elastic impedance inversion. Although the overall stratigraphic trend was preserved, the resulting Poisson's ratio section exhibited pronounced spatial smoothing. Thin coal seams were not sharply delineated, and tuning-dominated responses were significantly attenuated, reflecting the inherent limitations of linearized elastic impedance inversion when applied to thin-bed coal-bearing sequences.

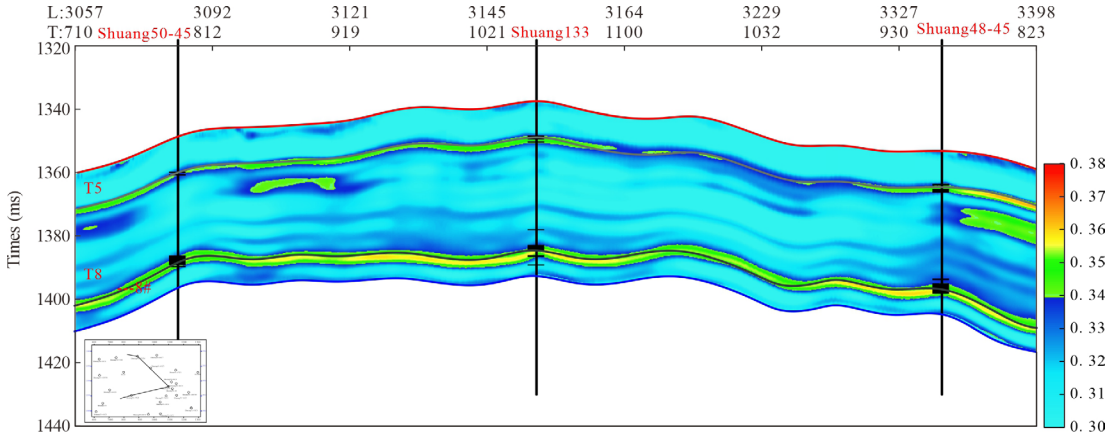



**Figure 12.** Inline profile of Poisson's ratio derived from single-branch convolutional neural network (SB-CNN) inversion. Although the overall Poisson's ratio range remained physically reasonable, the section showed localized lateral fluctuations and discontinuous features. The lack of physics-guided constraints led to reduced boundary coherence and limited the robustness of SB-CNN in resolving thin coal seams.

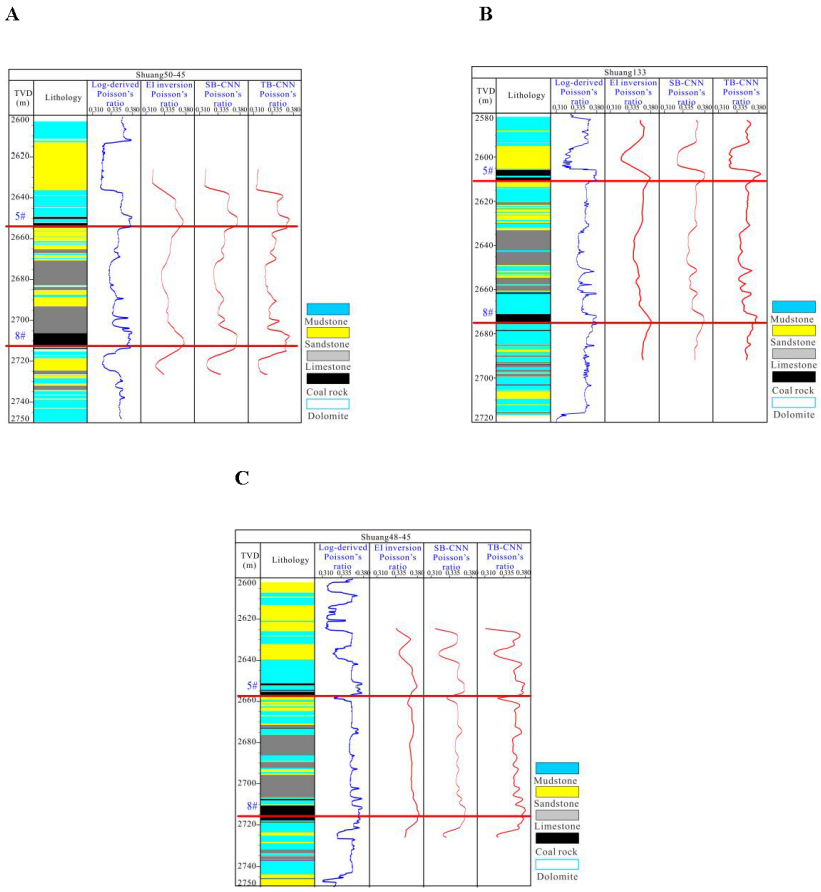



**Figure 13.** Comparison of Poisson's ratio profiles along three representative wells. (A) Shuang50–45, (B) Shuang133, and (C) Shuang48–45. Each well profile includes a lithology column, a log-derived Poisson's ratio computed from acoustic and density logs, and inversion results from elastic impedance (EI) inversion, single-branch convolutional neural network (SB-CNN), and the proposed two-branch (TB) CNN. The TB-CNN predictions (red curves) showed the highest consistency with the log-derived Poisson's ratio, particularly in reproducing the low Poisson's ratio coal seams and the moderate Poisson's ratio interbedded carbonate–clastic layers, and in delineating coal–seam boundaries and thin-bed variations, whereas the EI and SB-CNN results appeared smoother and less sensitive to thin interbeds.

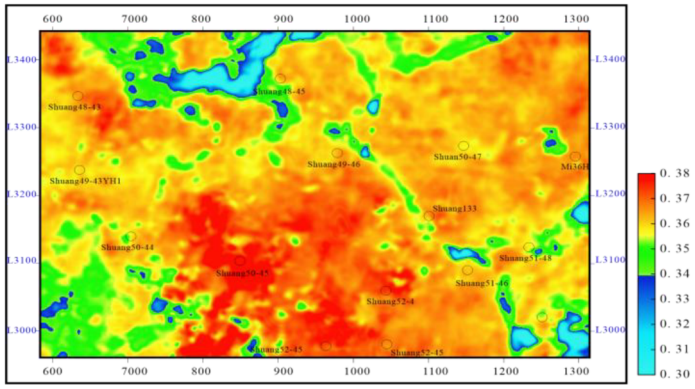



**Figure 14.** Horizon slice of Poisson's ratio predicted by a two-branch convolutional neural network along the 8# coal seam. Thin-seam variations and localized heterogeneity are clearly resolved.

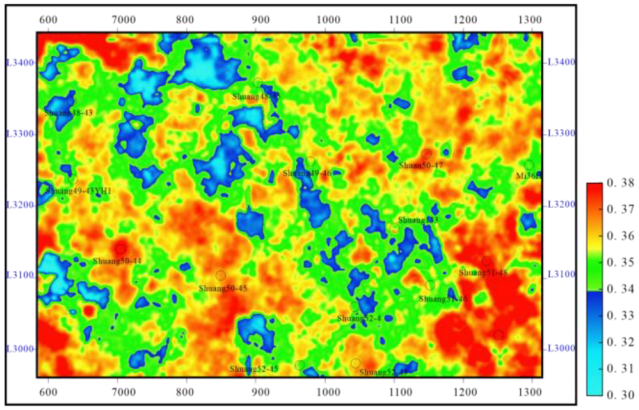



**Figure 15.** Horizon slice of Poisson's ratio derived from conventional EI inversion. Boundaries appear smoothed, and lateral variations are poorly captured

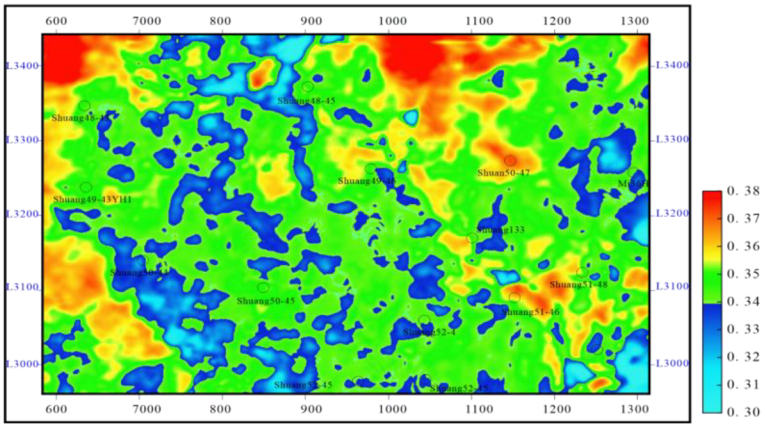



**Figure 16.** Horizon slice of Poisson's ratio derived from single-branch convolutional neural network inversion. Resolution is moderately improved compared with elastic impedance inversion, but lateral continuity and geological consistency remain limited.

seam boundaries, reflecting the inherent limitations of linearized inversion under thin-bed tuning and strong impedance interference. The SB-CNN result (Figure 16) showed moderate improvement in spatial resolution compared with EI inversion; however, discontinuous anomaly patterns and reduced geological coherence were still observed due to the absence of physics-guided constraints. Overall, the horizon-slice comparison demonstrated that the TB-CNN provided the most geologically consistent and spatially stable representation of ν along the 8# coal seam, achieving a superior balance between lateral resolution and continuity compared with both EI inversion and SB-CNN.

Three systematic advantages of the proposed TB-CNN were consistently observed from the comparative analyses: (i) boundary resolution: the upper and lower interfaces of coal seams were sharply delineated, significantly reducing vertical uncertainty induced by tuning effects; (ii) thickness sensitivity: gradual lateral thickening and thinning trends were preserved rather than flattened, as commonly observed in conventional EI inversion; (iii) Noise robustness: random seismic noise was effectively suppressed without excessive smoothing, while spurious local fluctuations were markedly reduced compared with the purely data-driven SB-CNN.

Overall, the TB-CNN effectively integrates the interpretability and stability of physics-guided inversion with the flexibility and resolution of data-driven learning. It provides a geologically consistent and physically reliable framework for ν inversion and CBM reservoir characterization, achieving consistent improvements in boundary sharpness, heterogeneity detection, and predictive robustness.

## 4. Discussion

This study presents a physics-constrained TB-CNN for direct ν inversion in thin coal-bearing gas reservoirs, addressing key limitations of conventional EI methods, including wavelet interference, bandwidth loss, and weak sensitivity to thin-bed tuning. By combining EI-derived Vp, Vs, and ρ with multi-angle seismic stacks, the proposed framework preserved large-scale geological consistency while enhancing thin-layer detectability through high-frequency seismic attributes. This dual-branch fusion strategy enabled clearer seam boundary delineation, improved thickness continuity, and more accurate tracking of localized thinning.

The two-branch architecture was central to the model's effectiveness. Branch 1 introduced physics-guided stability and provided rock-physics-consistent background trends, while Branch 2 focused on high-resolution reflectivity patterns sensitive to AVO behavior and thin-bed interference. Their latent fusion mitigated the limitations of linearized Zoeppritz approximations and reduced error propagation typically associated with sequential EI inversion. Direct inversion of ν avoided error propagation associated with sequential estimation of Vp and Vs, thereby improving prediction stability, while physical interpretability was maintained through rock-physics-guided regularization.

The robustness of the TB-CNN arises from both its data formulation and architectural design: (i) partial-angle stacking inherently suppressed random noise while retaining AVO-sensitive reflectivity, and (ii) the fusion of high-SNR, physics-consistent EI inputs with data-driven seismic features mitigated strong-reflection shielding at coal–rock interfaces. As a result, the model produced stable, high-resolution predictions even under low SNR conditions without excessive smoothing.

Unlike fully data-driven end-to-end approaches, the proposed TB-CNN leverages EI inversion as a physically meaningful foundation, enhancing transferability to geologically similar CBM blocks. By learning nonlinear residuals between EI-derived parameters and log-derived ν, the model corrects for angle inconsistencies, tuning effects, and local SNR deterioration. This EI-based feature space reduces dependency on the number of training wells and naturally supports transfer learning. In practice, the EI branch and early convolutional layers may be retained or frozen during domain adaptation, while only the fusion and regression layers are fine-tuned, making the approach practical for few-well scenarios. This feature is particularly valuable for CBM development in areas with sparse well control.

While the results are promising, some limitations remain. The model assumes relatively consistent wavelet characteristics across the study area, which may not hold in strongly heterogeneous or anisotropic environments. Additionally, the number of training wells remains limited, potentially affecting generalization in more complex lithological settings. Future work should explore anisotropy-aware rock-physics constraints,multi-scale feature fusion, and domain-adaptive transfer learning strategies.

## 5. Conclusion

This study introduced a practical and interpretable TB-CNN framework for ν inversion in thin coal-bearing gas reservoirs. The key findings are as follows:

(i) The proposed dual-branch architecture effectively fused EI-derived elastic attributes with multi-angle

seismic stacks, overcoming resolution loss, wavelet interference, and strong-reflection shielding typical of EI inversion in thin-bed and low SNR conditions.

(ii) The TB-CNN produced higher-resolution ν volumes with sharper seam boundaries, improved thin-layer detectability, and stronger agreement with well-log observations compared with conventional EI and single-branch CNN approaches.

(iii) Sensitivity and uncertainty analyses confirmed that the model is robust with respect to noise levels, tuning effects, and seam thickness, and maintains geological interpretability through rock-physics–guided regularization.

(iv) The hybrid EI–CNN design enhanced transferability, offering a scalable and adaptable framework suitable for few-well scenarios and lateral extension across geologically similar CBM blocks.

Overall, the TB-CNN provides a reliable tool for thin-bed elastic parameter inversion and CBM reservoir evaluation. Its ability to balance physics-guided stability with data-driven resolution makes it particularly promising for broader applications in unconventional reservoir characterization, including tight sandstone, shale, and carbonate formations.

## Acknowledgments

The authors gratefully acknowledge this support and thank all project collaborators for their assistance with data processing and interpretation.

## Funding

This study was supported by the commissioned project "Development of Seismic Prediction Methods for Coal and Rock Gas Reservoirs Based on Wave Theory" (Contract No. KFKT2024-29), funded by the Exploration and Development Research Institute of Changqing Oilfield Branch, CNPC, and Chengdu University of Technology.

## Conflict of interest

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

## Author contributions

*Conceptualization:* Fei Li, Mengbo Zhang, Qiang Liang
*Formal analysis:* Yongheng Zhang, Mengbo Zhang, Qingzhou Zhang
*Investigation:* Fei Li, Xiaojie Cui, Na Ni
*Methodology:* Mengbo Zhang, Qingzhou Zhang
*Visualization:* Xiaojie Cui ,Yongheng Zhang
*Writing–original draft:* Yongheng Zhang, Qiang Liang, Na Ni
*Writing–review & editing:* Fei Li, Mengbo Zhang, Qiang Liang

## Availability of data

Not applicable.

## References

1. Shuey RT. A simplification of the Zoeppritz equations. *Geophysics*. 1985;50(4):609-614.

   doi: 10.1190/1.1441936

2. Widess MB. How thin is a thin bed? *Geophysics*. 1973;38(6):1176-1180.

   doi: 10.1190/1.1440403

3. Ristow D, Rühl T. Fourier finite-difference migration. *Geophysics*. 1994;59(12):1882-1893.

   doi: 10.1190/1.1443575

4. Buland A, Omre H. Bayesian linearized AVO inversion. *Geophysics*. 2003;68(1):185-198.

   doi: 10.1190/1.1543206

5. Jardani A, Revil A, Slob E, Söllner W. Stochastic joint inversion of 2D seismic and seismoelectric signals in linear poroelastic materials: A numerical investigation. *Geophysics*. 2010;75(1):N19-N31.

   doi: 10.1190/1.3279833

6. Hampson DP, Schuelke JS, Quirein JA. Use of multiattribute transforms to predict log properties from seismic data. *Geophysics*. 2001;66(1):220-236.

   doi: 10.1190/1.1444899

7. Ge Q, Cao H, Yang Z, *et al*. High-resolution seismic impedance inversion integrating the closed-loop convolutional neural network and geostatistics: an application to the thin interbedded reservoir. *J Geophys Eng*. 2022;19(3):550-561.

   doi: 10.1093/jge/gxac035

8. Zhang B, Pu Y, Dai R, Cao D. Seismic poststack impedance inversion using geophysics-informed deep-learning neural network. *Interpretation*. 2025;13(2):T219-T232.

   doi: 10.1190/int-2024-0041.1

9. Chen Y, Fomel S. Random noise attenuation using local signal-and-noise orthogonalization. *Geophysics*. 2015;80(6):WD1-WD9.

   doi: 10.1190/geo2014-0227.1

10. Xue YJ, Zhang H, Zhang JQ, Wang XJ, Cao JX, Liu ZG. Local quantum filtering and denoising of seismic data. *Geophysics*. 2025;90(5):V455-V472.

doi: 10.1190/geo2024-0723.1

11. 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

12. Fang W, Fu L, Zhang M, Li Z. Seismic data interpolation based on U-net with texture loss. *Geophysics*. 2021;86(1):V41-V54.

doi: 10.1190/geo2019-0615.1

13. Zhu G, Chen X, Li J, Guo K. Data-Driven Seismic Impedance Inversion Based on Multi-Scale Strategy. *Remote Sens*. 2022;14(23):6056.

doi: 10.3390/rs14236056

14. Jiang X, Cao J, You C, Wang X, Du Z. Deep Neural Networks for Direct Hydrocarbon Detection in Prestack Seismic Data Based on AVO Analysis. *IEEE Trans Geosci Remote Sens*. 2023;61:1-9.

doi: 10.1109/tgrs.2023.3333910

15. Yang N, Li G, Zhao P, Zhang J, Zhao D. Porosity prediction from pre-stack seismic data via a data-driven approach. *J Appl Geophys*. 2023;211:104947.

doi: 10.1016/j.jappgeo.2023.104947

16. Rasht-Behesht M, Huber C, Shukla K, Karniadakis GE. Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions. *J Geophys Res Solid Earth*. 2022;127(5):e2021JB023120.

doi: 10.1029/2021jb023120

17. Sun J, Innanen K, Zhang T, Trad D. Implicit Seismic Full Waveform Inversion With Deep Neural Representation. *J Geophys Res Solid Earth*. 2023;128(3):e2022JB025964.

doi: 10.1029/2022jb025964

18. Jiang X, Cao J, Yang J, Liu J, Zhou P. AVO Analysis Combined With Teager-Kaiser Energy Methods for Hydrocarbon Detection. *IEEE Geosci Remote Sens Lett*. 2022;19:1.

doi: 10.1109/LGRS.2020.3029170

19. Xudong J, Junxing C, Jiangtao H, Xuanchen X, Jie L. Pre-stack gather optimization technology based on an improved bidimensional empirical mode decomposition method. *J Appl Geophys*. 2020;177:104026.

doi: 10.1016/j.jappgeo.2020.104026

20. Jiang X, Cao J, Zu S, Xu H, Wang J. Detection of hidden reservoirs under strong shielding based on bi-dimensional empirical mode decomposition and the Teager-Kaiser operator. *Geophys Prospect*. 2021;69(5):1086-1101.

doi: 10.1111/1365-2478.13073

21. Li P, Wang Y, Liu J, Liang M. Evaluation of Carbon Emission Efficiency and Analysis of Influencing Factors of Chinese Oil and Gas Enterprises. *Energy Sci Eng*. 2025;13(3):1156-1170.

doi: 10.1002/ese3.2055

22. Castagna JP, Sun S, Siegfried RW. Instantaneous spectral analysis: Detection of low-frequency shadows associated with hydrocarbons. *Lead Edge*. 2003;22(2):120-127.

doi: 10.1190/1.1559038

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Fracture prediction based on W-transform inversion spectral decomposition

**Xiuwei Wang[1,2,3], Ying Jia[2,3], Xilin Qin[1]*, Jun Wang[2,3], and Yinglong Li[1]**

[1]Key Laboratory of Oil and Gas Resources and Exploration Technology, Ministry of Education, Yangtze University, Wuhan, Hubei, China

[2]Research Institute of Exploration and Development, PetroChina Huabei Oilfield Company, Renqiu, Hebei, China

[3]CNPC International Chad Company, Beijing, China

## Abstract

Conventional ant-tracking technology is effective for fault identification, but its application is constrained by sensitivity to data quality and inadequate continuity of fault tracking results. This limitation creates an urgent need for optimization. To address this issue, this study proposes a frequency-divided ant-tracking method based on W-transform inversion spectral decomposition (WT-ISD) to improve the accuracy of fracture prediction in complex structural areas. By integrating the flexible time–frequency localization of the W-transform with sparse inversion theory, this method markedly improves the resolution and focusing performance of time–frequency spectra. This is achieved through the construction of an overcomplete dictionary and the imposition of L1 norm constraints. The proposed method not only provides a novel high-resolution time–frequency analysis tool for seismic signal processing but also establishes a theoretical basis for subsequent fine geological interpretation via its frequency-divided processing workflow. Taking the Raphia S Block in the Bongor Basin as a case study, this research first adopts WT-ISD to obtain high-resolution spectral decomposition results, generating single-frequency data volumes corresponding to different frequency components. Subsequently, structural smoothing filtering and boundary enhancement processing are applied to highlight discontinuity information. Then, the ant-tracking algorithm is introduced, combined with regional structural attitude constraints, to realize the identification of multi-scale fractures. Finally, red–green–blue attribute fusion technology is used to integrate responses from different frequencies and construct a fault distribution model. Practical application results indicate that this method not only improves the accuracy and spatial continuity of fracture prediction but also enhances its anti-noise capability. In particular, it exhibits excellent identification performance for large-, medium-, and small-scale fractures corresponding to 20 Hz, 35 Hz, and 50 Hz, respectively. This study verifies that the proposed method can provide reliable technical support for multi-scale fracture detection in complex structural areas and demonstrates important application value for fracture prediction in oil and gas exploration.

***Keywords:*** Frequency-divided ant-tracking; Fracture prediction; W-transform; Inversion spectral decomposition; Red–green–blue attribute fusion

***Corresponding author:**
Xilin Qin (qxl@yangtzeu.edu.cn)

**Citation:** Wang X, Jia Y, Qin X, Wang J, Li Y. Fracture prediction based on W-transform inversion spectral decomposition. *J Seismic Explor*. 2026;35(2):025010140. doi: 10.36922/JSE025010140

**Received:** December 30, 2025

**Revised:** February 6, 2026

**Accepted:** March 2, 2026

**Published online:** April 15, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Fractures serve as crucial reservoir spaces and seepage channels in oil and gas reservoirs. Their spatial distribution characteristics directly affect oil and gas accumulation and productivity, making them a key geological basis for formulating exploration and development plans.[1] Based on scale and genesis, fracture systems are generally classified into three categories. Large-scale faults and fault zones extend several kilometers and are typically formed by regional tectonic movements. Medium-scale structural fractures and fracture zones span tens to hundreds of meters and are mostly related to local structural deformation. Small-scale fractures such as microfractures and joints extend several meters to more than ten meters and are mainly controlled by diagenesis or local stress.[2] Different-scale fractures contribute differently to reservoir storage and seepage. Thus, the effective identification and detailed characterization of multi-scale fractures have long been a core challenge in the field of oil and gas geophysics.

Current fracture prediction methods are mainly divided into pre-stack and post-stack categories. Pre-stack methods, such as azimuthal anisotropy inversion, can evaluate fracture orientation and density using the directionality of seismic waves. Post-stack methods primarily rely on seismic attribute analysis, including discontinuity detection technologies such as coherence cubes, curvature attributes, and ant-tracking.[3,4] However, due to the multi-scale nature and genetic complexity of fracture systems, a single technology often fails to effectively characterize fracture networks of different scales simultaneously. Prediction results are usually a mixed reflection of multi-scale geophysical responses. This highlights the need for methods capable of separating and characterizing multi-scale fractures to improve the reliability and practicality of fracture prediction.

Seismic signals exhibit typical non-stationary characteristics. The traditional Fourier transform can only reflect the global frequency distribution of signals and cannot effectively characterize their local time–frequency evolution features. This limitation severely restricts the accurate extraction of detailed subsurface geological information. To address this issue, time–frequency analysis techniques have been widely applied in seismic data processing and interpretation, and are mainly categorized into linear and nonlinear types. Linear time–frequency analyses include the short-time Fourier transform (STFT),[5] wavelet transform,[6] and S-transform (ST).[7] Nonlinear time–frequency analyses encompass the Wigner–Ville distribution,[8] matching pursuit,[9] and frequency reassignment.[10] These techniques offer solid technical support for improving the identification accuracy of thin layers and complex geological bodies.

Specifically, the spectral decomposition method proposed by Partyka *et al.*[11] has successfully revealed subtle geological details that are indistinguishable in conventional broadband data by generating discrete single-frequency data volumes. Furthermore, Zeng *et al.*[12] confirmed that single-frequency data volumes possess unique advantages in characterizing the spatial distribution of geological anomalies, providing a new technical approach for the detailed characterization of oil and gas reservoirs.

Nevertheless, traditional linear time–frequency analyses are constrained by the inherent properties of window functions, generally suffering from insufficient resolution in low-frequency regions. This directly leads to inadequate accuracy in identifying key geological features of oil and gas reservoirs, which has become a core bottleneck restricting the efficiency of oil and gas exploration. To address this technical challenge, Wang[13] proposed the W-transform (WT). Its core innovation lies in optimizing the window kernel function of traditional linear transforms using the instantaneous dominant frequency of seismic data and in constructing an improved dominant frequency estimation formula. These improvements not only overcome the theoretical limitations of traditional linear time–frequency analysis but also significantly enhance its practical application value in production scenarios such as reservoir description.

Since the proposal of the WT, researchers have conducted extensive targeted studies on the optimization and upgrading of its window kernel function, leading to the development of several modified versions with important application value. Li *et al.*[14] conducted an in-depth analysis of the mathematical model of the conventional WT and discovered a derivative singularity near the dominant frequency, which causes distortion in the extraction of time–frequency features in the dominant frequency region. To resolve this defect, they reconstructed the window kernel function into a polynomial form and proposed the generalized WT, effectively improving the time–frequency resolution of the WT near the dominant frequency.

The three-parameter WT proposed by Chen *et al.*[15] shares the same design philosophy as the generalized WT. By introducing three auxiliary control parameters, it achieves precise adaptation of the time–frequency resolution of the smoothed Gaussian window, further enhancing the characterization accuracy of the time–frequency spectra of seismic data. Luo *et al.*[16] integrated the core idea of the synchrosqueezing transform into the WT framework and proposed the synchrosqueezing WT. This significantly improves the performance of the WT in detecting small-scale geological bodies such as thin layers

and channels, offering more powerful technical support for the detailed characterization of small-scale geological bodies.

In addition, the normalized WT proposed by Chen *et al.*[17] effectively solves the spectral energy splitting problem of the WT in broadband seismic signal analysis. It achieves this by normalizing the standard deviation along the time-varying dominant frequency, significantly improving the time–frequency resolution and reservoir interpretation accuracy. These modified versions continuously enrich and improve the theoretical system and application scenarios of the WT, providing more efficient technical approaches for oil and gas exploration under complex geological conditions.

Ant-tracking technology provides another effective approach for fracture detection. Originating from the ant colony algorithm proposed by Dorigo *et al.*[18], its global optimization characteristics lay a theoretical foundation for seismic discontinuity identification. In recent years, multi-scale ant-tracking technology has promoted the refined development of fracture characterization. For example, Yuan and Simons[19] proposed a multi-scale full-waveform inversion method based on the wavelet transform. It achieves step-by-step inversion from long wavelengths to seismic phase details through wavelet decomposition, significantly improving convergence and stability. Wen *et al.*[20] proposed and applied a fracture identification method based on the multi-channel synchrosqueezing generalized ST. By leveraging its high time–frequency focusing capability, it maximizes the retention of fracture information and reduces errors in time–frequency conversion, realizing multi-scale fault characterization.

Moreover, Liu *et al.*[21] proposed a novel time–frequency analysis method called the synchrosqueezing extraction transform. It aims to improve the concentration of the time–frequency representation of non-stationary vibration signals, effectively extracting time-varying information from the signals and obtaining clearer time–frequency representations. Xie *et al.* established a frequency-divided optimized ant-tracking workflow.[22] Through multi-frequency band collaborative processing, it significantly enhances the continuity of fracture prediction and the identification ability of small-scale fractures.

Meanwhile, time–frequency analysis methods are continuously evolving. For instance, the improved synchrosqueezing transform proposed by Yu *et al.*,[23] the time-synchroextracting generalized ST extraction method established by Zhang *et al.*,[24] the matched demodulation synchrosqueezing ST designed by Wang *et al.*,[25] and the short-time fractional Fourier transform spectrum proposed by Tian[26] have all improved time–frequency focusing and reservoir characterization accuracy from different perspectives.

However, existing methods still have limitations in separating multi-scale fracture responses. How to effectively perform high-resolution time–frequency analysis for fracture detection algorithms remains a topic worthy of further investigation. Therefore, the present study integrates the advantages of sparse inversion and the WT to construct a novel high-resolution time–frequency analysis method—WT inversion spectral decomposition (WT-ISD). It also combines WT-ISD with ant-tracking technology to propose a multi-scale fracture prediction method based on frequency-divided processing. This method aims to obtain single-frequency data volumes that are sensitive to fractures of specific scales through spectral decomposition. Additionally, the study uses ant-tracking under regional structural constraints to realize the separation and detection of fractures at different scales. Finally, it constructs a complete fault distribution model through red–green–blue (RGB) attribute fusion technology. This study verifies the effectiveness of the method in improving fracture prediction accuracy, continuity, and anti-noise capability using data from an actual study area, providing reliable technical support for multi-scale fracture characterization in complex structural areas.

## 2. Methodology

The frequency-divided ant-tracking method based on WT-ISD offers distinct advantages in fracture interpretation, with the specific workflow illustrated in Figure 1. Based on post-stack three-dimensional (3D) seismic data, spectral analysis of the seismic data was first conducted to clarify the dominant frequency band range of the dataset. Subsequently, by virtue of the high-precision time–frequency analysis method WT-ISD, 12 single-frequency seismic data volumes with different dominant frequencies were extracted within the effective frequency band. These 12 single-frequency volumes were then preprocessed to suppress noise and improve the signal-to-noise ratio, aiming to enhance the discontinuity characteristics in the data. Then, an in-depth analysis was conducted on the detection response differences of the post-stack seismic variance attribute among single-frequency volumes of different frequency bands.

Fracture information corresponding to each dominant frequency was extracted through discontinuity detection, and through comprehensive comparison, three single-frequency attribute data volumes from different frequency bands that best highlight feature characteristics at different scales and are suitable for ant-tracking–based fracture

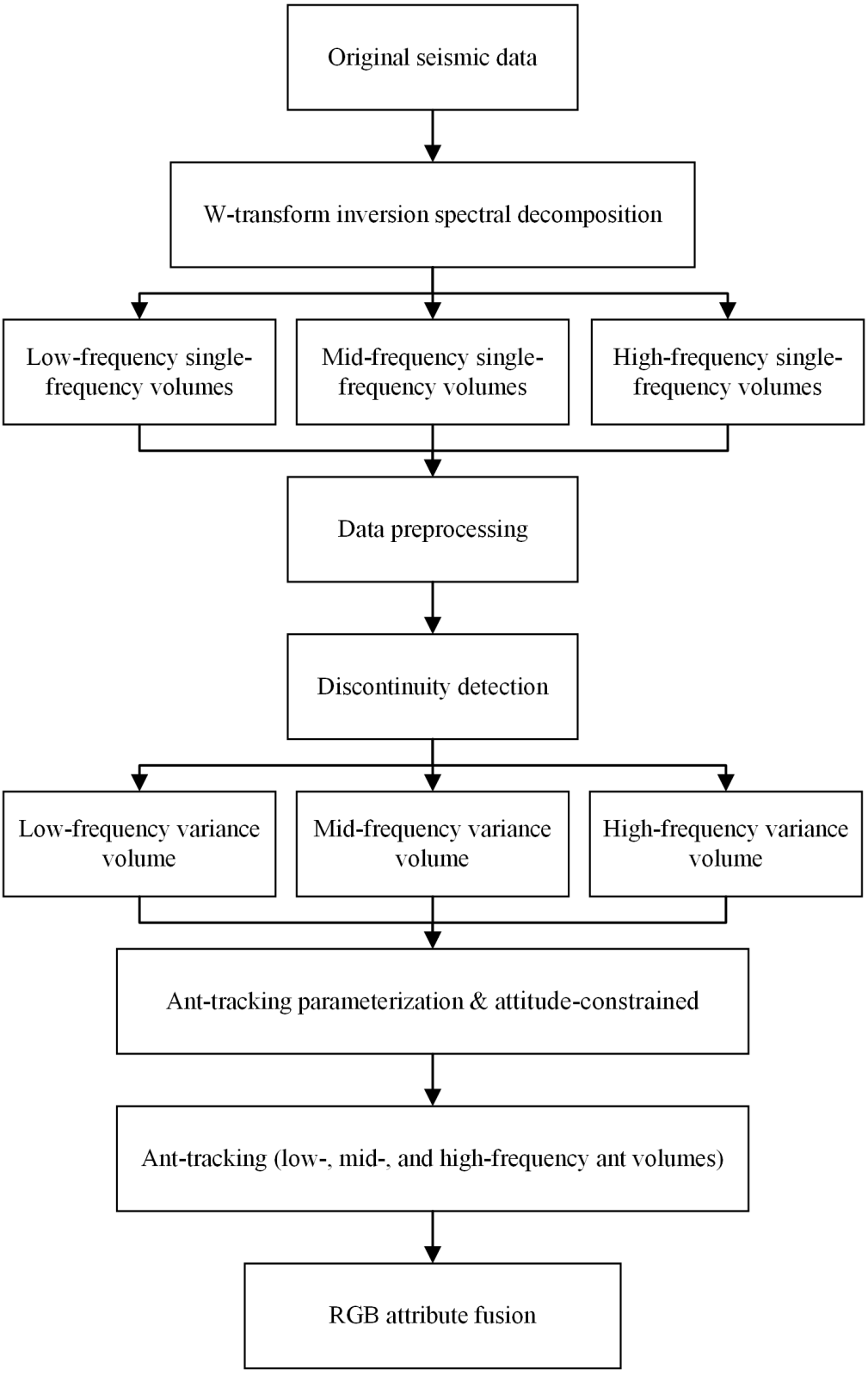



**Figure 1.** Workflow of frequency-divided ant-tracking
Abbreviation: RGB: Red–green–blue.

detection were selected. Finally, when applying the ant-tracking method to conduct fracture tracing on the optimally selected single-frequency variance data volumes that can characterize fractures at three different scales, attitude control must be implemented in combination with the regional structural background to ensure that the tracing results are consistent with geological reality, thereby realizing the effective prediction and detailed characterization of fractures of different scales.

Ultimately, results were generated and verified through attribute fusion. The RGB attribute fusion technology integrates single-frequency data volumes with different frequency components, enabling comprehensive consideration of the multi-scale development characteristics of fractures under each frequency band and effectively improving prediction accuracy. Through this workflow, the frequency-divided ant-tracking fracture prediction method based on WT-ISD can significantly enhance the accuracy and reliability of fault interpretation.

### 2.1. W-transform inversion spectral decomposition

Time–frequency analysis technology is a class of mathematical methods capable of synchronously revealing signal characteristics in both the time and frequency domains, which is of great significance for processing non-stationary signals. As a typical non-stationary signal, seismic signals have frequency components that change with time and contain abundant geological information.[27] The traditional Fourier transform can only provide the global frequency distribution of signals and cannot characterize the local frequency evolution characteristics in the time dimension; therefore, time–frequency analysis methods play an irreplaceable role in seismic data interpretation and reservoir prediction.

The core of linear time–frequency analysis methods lies in constructing a time–frequency localized window function and characterizing the energy distribution of signals on the time–frequency plane through convolution operations between the window function and the signal. Specifically, such methods extract local features of signals at different times and frequencies through convolution operations between signals and time-window kernel functions, thereby realizing the time–frequency spectral characterization of non-stationary signals.

To further improve the resolution and stability of time–frequency analysis, inversion spectral decomposition methods introduce inversion concepts and sparse constraint regularization into the time–frequency analysis framework. Qin *et al.*[28] utilized the advantage of the generalized ST in flexibly adjusting the time–frequency window to construct a spectral decomposition inversion model based on sparse constraints, significantly enhancing the energy concentration and resolution of time–frequency spectra. The key to this method is to reformulate the traditional linear transform process as an underdetermined linear inverse problem and obtain the optimal solution with clear physical significance by introducing the sparsity of signals in the time–frequency domain as a prior constraint.

In this study, the WT with adjustable time–frequency window characteristics is introduced to establish an inversion spectral decomposition model based on the WT. WT-ISD not only inherits the advantages of the sparse inversion framework in improving time–frequency resolution but also provides clearer physical significance and more flexible adjustment mechanisms for window function parameters. It can better adapt to the time-varying behavior of different frequency-band components in seismic signals by adjusting its time–frequency focusing characteristics, thereby providing a new technical approach for fine interpretation tasks such as fracture detection and thin-layer identification. The WT of a continuous-time signal $x(t)$ is defined as follows:

$$X(\tau,\ f)=\int_{-\infty}^{\infty}x(t)w(t-\tau;f)e^{-i2\pi ft}dt \tag{1}$$

where $X(\tau,f)$ is the time–frequency spectrum, $\tau$ is the time-shift parameter, and $f$ is the frequency. The $w(t;f)$ employs an adjustable Gaussian window:

$$w(t;\ f)=\frac{\lambda f}{\sqrt{2\pi}}e^{-\frac{t^2\lambda^2f^2}{2}} \tag{2}$$

where $\lambda$ is the scale factor, which is the core parameter balancing time–frequency resolution. An increase in $\lambda$ narrows the time window, improving time resolution but reducing frequency resolution; conversely, a decrease in $\lambda$ improves frequency resolution. The window function of the WT achieves continuous adjustment of time–frequency localization characteristics through a single parameter $\lambda$, and compared with the traditional ST, it has greater flexibility and adaptability.

The inverse WT ensures that signals can be completely reconstructed from their time–frequency spectra, and its expression reveals that signals can be represented as a linear combination of time–frequency atoms:

$$x(t)=\iint X(\tau,\ f)w(t-\tau,f)e^{i2\pi ft}dfd\tau \tag{3}$$

The above inverse transform can also be explicitly expressed in the form of a convolution:

$$x(t)=\int\left[X(\tau,f)\otimes w(t,f)\right]e^{i2\pi ft}dt \tag{4}$$

where $\otimes$ denotes the convolution operation. This expression provides a direct theoretical basis for constructing discrete inversion models. For numerical calculation, the continuous inverse transform formula is discretized. Let the discrete seismic signal be the vector $x=\begin{bmatrix}x_{t_1} & x_{t_2} & \cdots & x_{t_N}\end{bmatrix}^T$, and the target frequency set be $\{f_1 \quad f_2 \quad \cdots \quad f_M\}$; the following discrete inversion model can be constructed:

$$\begin{bmatrix}x_{t1}\\ x_{t2}\\ \vdots\\ x_{tN}\end{bmatrix}=\begin{bmatrix}G_{f1} & G_{f2} & \cdots & G_{fM}\end{bmatrix}\begin{bmatrix}m_{t_1,f_1}e^{i2\pi f_1 t_1}\\ \vdots\\ m_{t_N,f_1}e^{i2\pi f_1 t_N}\\ \vdots\\ m_{t_1,f_M}e^{i2\pi f_M t_1}\\ \vdots\\ m_{t_N,f_M}e^{i2\pi f_M t_N}\end{bmatrix} \tag{5}$$

where $G_{f_i}$ is the convolution matrix composed of the Gaussian window function $w(t;f_i)$ corresponding to frequency $f_i$, and $m$ is the vectorized form of the time-frequency coefficient matrix to be solved. **Equation 5** can be simplified as:

$$x=Gm \tag{6}$$

In seismic data analysis, $x$ represents seismic records, $G$ is the overcomplete dictionary matrix composed of WT window functions, and $m$ is the high-resolution time–frequency spectrum to be solved. Since the reflection coefficient sequence usually satisfies the sparse distribution assumption, its time–frequency spectrum should also be sparse in the time–frequency domain, which provides a physical basis for introducing sparse constraints. Based on the prior sparsity of seismic signals in the time–frequency domain, the L1-norm sparse constraint is introduced to construct the following optimization problem:

$$\hat{m}=\arg\min_m\{\|Gm-x\|_2^2+\beta\|m\|_1\} \tag{7}$$

where $\beta$ is the regularization factor used to balance the weight between the data fitting term $\|Gm-x\|_2^2$ and the sparse constraint term $\|m\|_1$. An appropriate value of $\beta$ can be determined through the L-curve method. This optimization problem is efficiently solved using the spectral projected gradient L1 minimization algorithm (SPGL1) to obtain the high-resolution time–frequency spectrum $\hat{m}$.

Based on the solved high-resolution time–frequency spectrum $\hat{m}$, high-fidelity frequency-divided processing can be realized under the inversion framework. For the target center frequency $f_c$, a Gaussian weight function $W(f;f_c,\sigma)$ is designed to weight the time–frequency spectrum:

$$m_{f_c}=\hat{m}\circ W(f;f_c,\sigma) \tag{8}$$

where $\circ$ denotes the element-wise multiplication (Hadamard product) along the frequency dimension, and $\sigma$ is the parameter controlling the frequency-divided bandwidth. The weighted time–frequency coefficient $m_{f_c}$ is reconstructed through the inverse WT model to obtain the frequency-divided seismic data $x_{f_c}$ corresponding to the frequency band:

$$x_{f_c}=Gm_{f_c} \tag{9}$$

This inversion and weighted reconstruction process is repeated for each trace in the 3D seismic data volume, and finally, a series of high-resolution frequency-divided data volumes reflecting the reflection characteristics of different frequency bands are obtained.

### 2.2. Spectral analysis performance

To systematically evaluate the performance of WT-ISD, a comparative analysis between WT-ISD and other time–frequency analysis methods was conducted. These methods include STFT, ST, and the WT. The comparison used both synthetic signals and actual seismic signals.

Firstly, comparative verification was performed based on synthetic signals. The synthetic signal was composed of three Ricker wavelets with dominant frequencies of 20 Hz, 35 Hz, and 50Hz. Their time positions were set at 0.25 s, 0.55 s, and 0.85 s, respectively. The comparative results, as shown in Figure 2, clearly illustrate the characteristics of each method. Limited by the fixed time window, STFT has acceptable resolution for the 20 Hz low-frequency wavelet. However, it exhibits significant energy diffusion and side lobe interference during time–frequency decomposition of the 35 Hz and 50 Hz wavelets, resulting in poor time–frequency focusing. ST can reflect the energy distribution of each frequency component but has insufficient time resolution in the low-frequency region, leading to blurred energy clusters.

In contrast, the WT introduces the instantaneous frequency parameter and a dynamic scale adjustment mechanism. It provides higher time resolution in the 20 Hz low-frequency band and enhances frequency resolution in the 50 Hz high-frequency band, thereby achieving overall optimization of time–frequency resolution. However, for strongly non-stationary signals with rapidly changing instantaneous frequencies (e.g., seismic data), the WT still suffers from energy diffusion due to its fixed parameterized window function. To address this limitation, WT-ISD realizes adaptive resolution optimization through sparse constraints. The results show that WT-ISD can significantly improve the focusing of time–frequency energy.

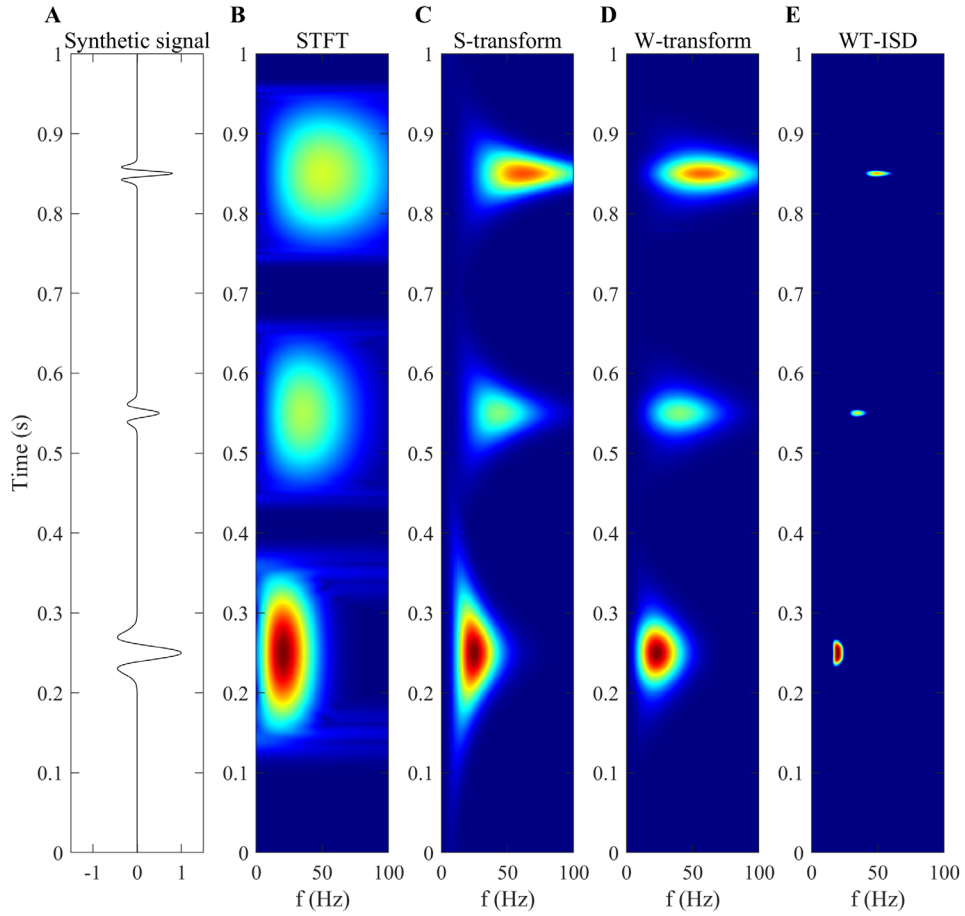



**Figure 2.** Spectrum of synthetic signal with different time–frequency analysis methods. (A) Synthetic signal. (B) STFT. (C) S-transform. (D) W-transform. (E) WT-ISD.
Abbreviations: STFT: Short-time Fourier transform; WT-ISD: W-transform inversion spectral decomposition.

To further verify the applicability of the method to actual seismic signals, a single-trace of actual seismic data from the study area was selected for testing. Figure 3 shows that WT-ISD also exhibits excellent performance when processing non-stationary and non-ideal actual seismic waveforms. The obtained time–frequency spectrum not only has the highest energy concentration, which enables finer characterization of the frequency evolution of effective signals over time, but also demonstrates stronger robustness against random noise, facilitating the identification and extraction of weak signals. The STFT and ST spectra show noticeable energy blurring, and although the WT effect is improved, its time–frequency localization details remain inferior to those of WT-ISD. Overall, WT-ISD outperforms other time–frequency analysis methods in both time–frequency resolution and noise suppression capability, providing a solid foundation for subsequent high-precision spectral decomposition and effective extraction of frequency response characteristics of fractures of different scales.

## 3. Application examples

### 3.1. Regional geological overview

The Bongor Basin extends overall in a Northeast (NE)–Southeast (SE) direction, exhibiting a typical spindle-shaped geometry. It reaches a maximum width of 80 km at its central part and gradually tapers toward both ends, with a total area of approximately $1.8 \times 10^4$ km$^2$. The Raphia S block, the focus area of this study, is located in the northern slope belt of the basin, where the strata show a prominent wedge-shaped thinning trend toward the north. The northern slope belt hosts a multi-stage, genetically complex fault system, which can be divided into two categories based on their timing of formation: basement interior faults and basin-bounding faults. Basement interior faults

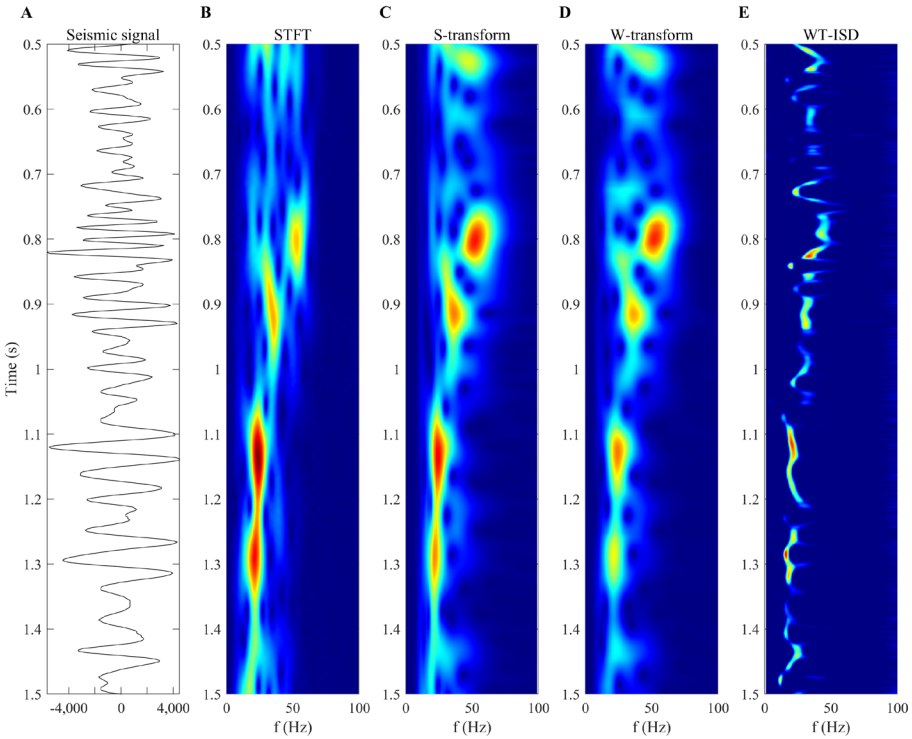



**Figure 3.** Spectrum of single channel seismic signal with different time–frequency analysis methods. (A) Seismic signal. (B) STFT. (C) S-transform. (D) W-transform. (E) WT-ISD.
Abbreviations: STFT: Short-time Fourier transform; WT-ISD: W-transform inversion spectral decomposition.

formed during the pre-rift stage of basin evolution and primarily control the internal structural differentiation of buried hills. Basin-bounding faults developed during the syn-rift and post-rift stages and dominate the structural evolution of the sags.

The basin is subdivided from south to north into four first-order tectonic units: the Southern Depression, Southern Uplift, Central Depression, and Northern Slope Belt (Figure 4). The Raphia buried hill belt, characterized by distinct structural features, is situated within the

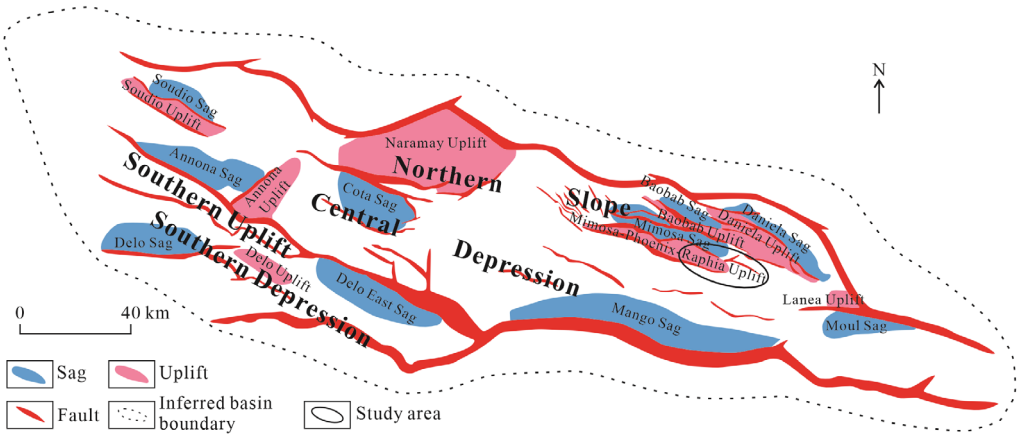



**Figure 4.** Tectonic units of bongor basin. Redrawn based on the data from yu *et al.*[29]

Northern Slope Belt.

### 3.2. Spectral decomposition

Spectral decomposition techniques map seismic signals from the time domain to the frequency domain. By analyzing the evolutionary characteristics of amplitude and phase with frequency, they can effectively identify key geological information—including reservoir distribution, physical properties, and hydrocarbon potential.[30] This technique suppresses the mutual interference among different frequency components in the time domain, thereby delivering geological interpretations with higher resolution than conventional methods.

In this study, the WT-ISD method was employed for the time–frequency decomposition of seismic data. Its primary objective was to extract single-frequency data volumes corresponding to distinct frequency components, and the procedure consisted of a systematic sparse inversion and signal reconstruction workflow.

The method starts with trace-by-trace preprocessing of raw seismic data. This step includes obtaining analytic signals via the Hilbert transform and constructing an overcomplete dictionary matrix composed of frequency-dependent Gaussian window functions according to the target frequency band. The core innovation lies in reformulating the conventional WT as a basis pursuit denoising inverse problem. An L1-norm was introduced as a sparse constraint, and the optimal high-resolution time–frequency coefficient matrix was solved using SPGL1. Subsequently, within the inversion framework, a Gaussian frequency-weighting function centered at the target dominant frequency was designed. This function performs spectral focusing weighting on the time–frequency coefficients. The time-domain signals of the target frequency band were then accurately reconstructed using the adjoint of the dictionary matrix. Finally, the inversion and weighted reconstruction processes were implemented independently for each trace in the 3D seismic volume. A

**Algorithm 1. W-transform inversion spectral decomposition**

Inputs: 3D post-stack seismic data volume $X \in \mathbb{R}^{N_x \times N_y \times N_t}$, time sampling interval $\Delta t$, target frequency range $[f_{\min}, f_{\max}]$ with sampling $\Delta f$, scale factor $\lambda$, sparsity constraint parameter $\sigma$
Outputs: Frequency-divided seismic data volumes $\{X_{f_c}\}$ for each center frequency $f_c$
1: //Step 1: Construct overcomplete W-transform dictionary
2: Initialize frequency set $\mathcal{F} \leftarrow \{f_{\min}, f_{\min} + \Delta f, \ldots, f_{\max}\}$
3: for each frequency $f_j \in \mathcal{F}$ do
4: Generate frequency-dependent Gaussian window: $w_j(t) = \frac{\lambda f_j}{\sqrt{2\pi}} \exp\left(-\frac{t^2 \lambda^2 f_j^2}{2}\right)$
5: Precompute window FFT: $\mathbf{W}_j, \mathbf{W}_j^{\text{flip}} \leftarrow \text{FFT}(w_j), \text{FFT}(\text{flip}(w_j))$
6: end for
7: Construct phase modulation matrix $\mathbf{E}$ with $\mathbf{E}_{t,j} = e^{-i2\pi f_j t}$
8: //Step 2: Trace-wise sparse inversion (parallelizable)
9: for each seismic trace $\mathbf{x} \in \mathbf{X}$(dimension $N_t \times 1$) do
10: Compute analytic signal: $\mathbf{x}_a \leftarrow$ Hibert $(\mathbf{x})$
11: Define linear operator $\mathcal{A}(\mathbf{m}) = \mathbf{G}\mathbf{m}$ ,where $\mathbf{G}$ encodes convolution with $\mathbf{W}_j$ and modulation with $\mathbf{E}$
12: Solve basis pursuit denoising problem:
13: $\widehat{\mathbf{m}} = \arg\min_{\mathbf{m}} \| \mathbf{m} \|_1$ s.t. $\mathcal{A}(\mathbf{m}) - \mathbf{x}_a \|_2 \le \sigma$
14: Obtain time-frequency coefficient matrix $\mathbf{M} \in \mathbb{R}^{N_t \times |\mathcal{I}|}$ via SPGL1 algorithm
15: Store $\mathbf{M}$ to $\mathbf{M}_{\text{all}}$
16: end for
17: //Step 3: Frequency band extraction and reconstruction
18: for each target center frequency $f_c$ do
19: Construct Gaussian weight vector: $\mathbf{v}_c(f) = \exp\left(-\frac{f^2}{2f_c^2}\right)$
20: for each time–frequency matrix $\mathbf{M}$ do
21: Apply frequency weighting: $\mathbf{M}_c \leftarrow \mathbf{M} \odot \mathbf{v}_c$
22: Reconstruct single-frequency signal: $\mathbf{x}_{f_c} \leftarrow \text{Real}\left(\mathcal{A}^*(\mathbf{M}_c)\right)$
23: end for
24: Combine all $\mathbf{x}_{f_c}$ to form frequency-divided volume $\mathbf{X}_{f_c}$
25: end for
26: return $\{\mathbf{X}_{f_c}\}$
Key parameters:
$\lambda$ : Scale factor balancing time–frequency resolution (typically 1.0–4.0)
$\sigma$ : Inversion tolerance controlling sparsity level
$\beta$ : Gaussian bandwidth controlling frequency division sharpness

Abbreviations: 3D: Three-dimensional; FFT: Fast Fourier transform; SPGL1: Spectral projected gradient for L1 minimization algorithm.

series of high-resolution frequency-divided data volumes with concentrated energy and clear boundaries are efficiently generated via parallel computing.

Notably, the WT-ISD method incorporates an L1-norm sparse constraint. It transforms conventional linear time–frequency analysis into a model-driven nonlinear inversion problem, whose solution relies on the iterative optimization of the SPGL1 algorithm. As a result, its computational load is significantly higher than that of traditional closed-form time–frequency transforms.

The major computational overhead stems from two sources. The first is the construction of the frequency dictionary and convolution kernels, along with the subsequent convolution operations. The second is the multiple iterations of sparse inversion and repeated applications of linear operators. The computational complexity increases approximately linearly with the frequency sampling density and the length of the time window, and is further modulated by the number of iterations and convergence tolerance. To mitigate this computational burden, a single inversion, multi-frequency reconstruction strategy was adopted in this study. Only one SPGL1-based sparse inversion was performed per trace to obtain the time–frequency coefficients. All subsequent dominant-frequency data volumes are generated from these coefficients through efficient convolution and linear combinations of frequency-band weights. This design avoids repeated L1-norm iterations for each individual frequency band.

Meanwhile, the fast Fourier transform of frequency-dependent Gaussian windows and the dominant-frequency weight matrix were precomputed and reused at the beginning of the experiment. This measure substantially reduces the overhead of inner-loop operations. Trace-by-trace processing was parallelized across a multi-core platform. Additionally, the maximum number of iterations and convergence tolerance were constrained to control computational costs. This implementation achieves acceptable overall computational efficiency while preserving the advantage of high time–frequency resolution, meeting the demands of large-scale seismic volume processing.

In essence, WT-ISD upgrades linear time–frequency analysis to a model-driven sparse inversion framework. It effectively overcomes the inherent limitations on time–frequency resolution imposed by the Heisenberg uncertainty principle. Accordingly, it provides a reliable foundation of time–frequency features for spectrum-based fine reservoir characterization.

In practical implementation, WT-ISD is directly applied to the full-area 3D seismic data. Spectral analysis is first conducted to determine the dominant frequency band of the dataset. Using this high-precision time–frequency method, 12 single-frequency seismic data volumes with distinctly dominant frequencies are extracted within the effective frequency band.

To systematically evaluate the response relationship between frequency and fracture scale, 12 characteristic frequency-divided data volumes covering 5–60 Hz were selected. By comparing the spatial responses of different frequency volumes to structural discontinuities and stratigraphic attributes, the indicative relationships between frequency and fracture development scale were established. The findings revealed that low-frequency volumes show strong responses to the outlines of large-scale faults, medium-frequency volumes achieve optimal performance in identifying medium-scale fracture zones, and high-frequency volumes exhibit high sensitivity to microfractures and subtle geological features.

Figure 5 presents the original seismic profile and the spectral analysis results of the raw seismic data. Additionally, Figure 6 displays the seismic profiles of the 12 frequency-divided data volumes with different dominant frequencies after spectral decomposition. The results demonstrate that each single-frequency volume highlights structural and stratigraphic information of different hierarchical levels. Low-frequency volumes clearly delineate the planar distribution of large-scale faults. Medium-frequency volumes reveal more abundant internal structural details while retaining the outlines of major faults. High-frequency volumes further enhance the response signals of minor faults and small-scale fractures. These frequency-divided data volumes effectively separate the seismic response characteristics of geological bodies at different scales. They provide high-quality input data for subsequent discontinuity detection and multi-scale fracture identification using frequency-divided ant-tracking method.

### 3.3. Data preprocessing

Structure-oriented filtering technology, a key method in seismic data processing, can effectively utilize the dip angle and azimuth angle information of geological targets, significantly improving the accuracy of attribute prediction and the detection efficiency of target bodies. While enhancing the signal-to-noise ratio of the seismic data volume and preserving the basic characteristics of the original seismic signals, this technology can specifically strengthen the discontinuity of seismic events, generate clearer imaging of faults and fractures, and provide robust support for the accurate identification of micro-faults.[31]

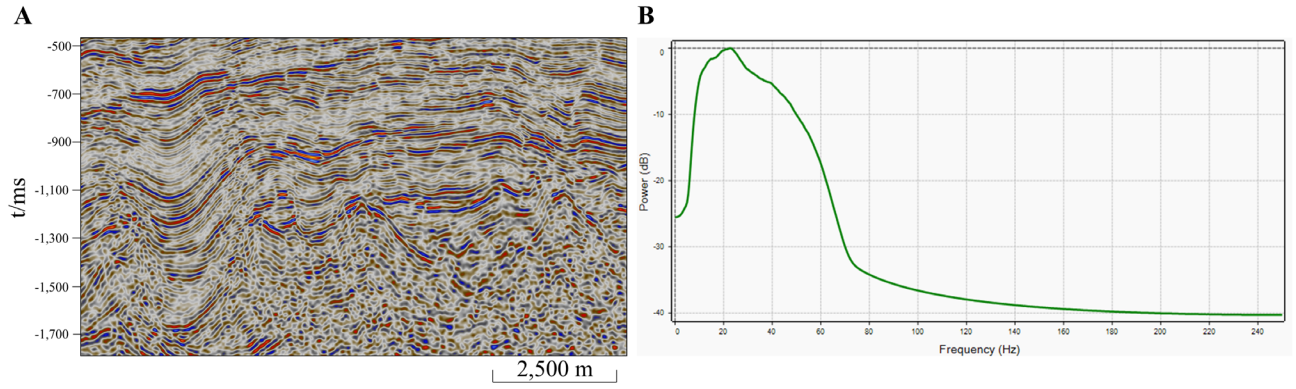



**Figure 5.** Original seismic profile and spectral analysis results of the raw seismic data in the Raphia S Block, Bongor Basin. (A) Post-stack seismic profile (scale bar = 2,500 m) and (B) amplitude spectrum analysis result of the raw seismic data, showing the power distribution variation with frequency.

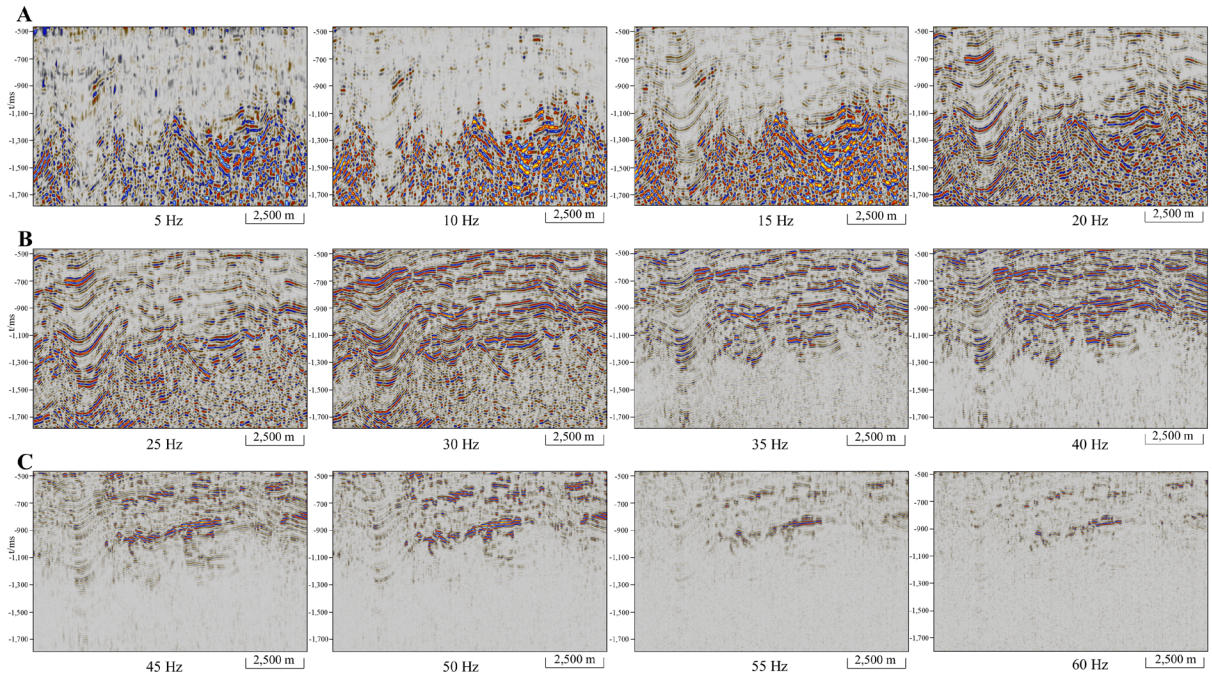



**Figure 6.** Single-frequency seismic profiles (scale bar = 2,500 m): (A) Low-, (B) medium-, and (C) high-frequency bands

Figure 7 shows the processing effects of frequency-divided ant volumes in different frequency bands after structure-oriented filtering. By comparing the results before and after filtering, it can be observed that the random noise in the original seismic data has been effectively suppressed, the continuity of seismic events is significantly improved, and the fault boundary characteristics within the dashed box are more prominent, thereby laying a solid foundation for the subsequent detailed characterization of fracture systems.

### 3.4. Discontinuity detection

After structure-oriented filtering, the resolution and clarity of faults have been improved to a certain extent, but some faults still suffer from issues such as blurred boundaries and insufficient spatial continuity. To address this limitation, discontinuity detection technology was further adopted to enhance the spatial continuity of fractures and highlight fault boundary characteristics.

Currently, there are several commonly used discontinuity detection methods, whose applicability and performance vary significantly depending on the geological conditions of different study areas. Among them, the variance volume attribute, as a classic post-stack seismic attribute, offers advantages of wide applicability and stable identification performance, making it a commonly used

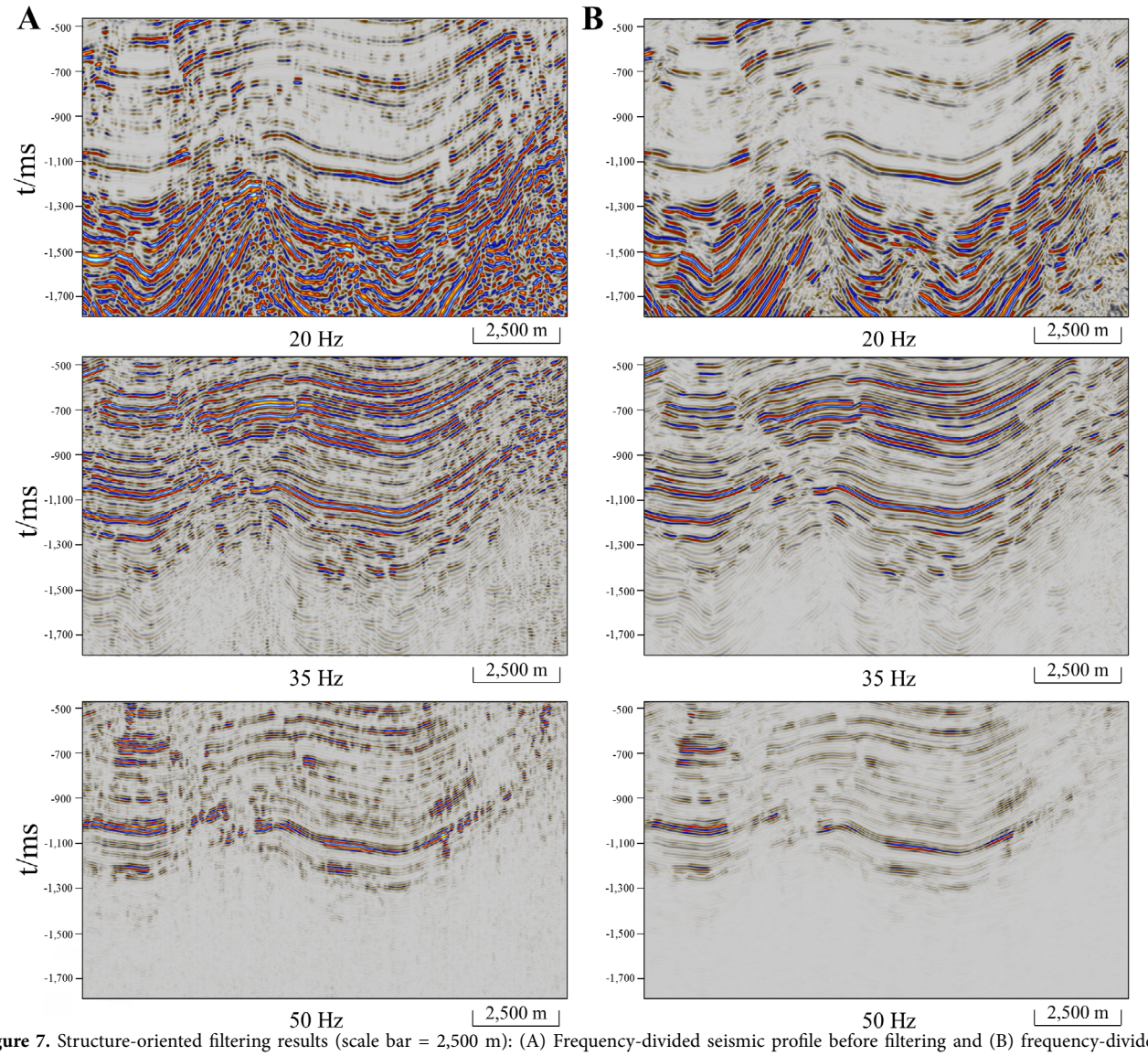



**Figure 7.** Structure-oriented filtering results (scale bar = 2,500 m): (A) Frequency-divided seismic profile before filtering and (B) frequency-divided seismic profile after structure-oriented filtering

and effective method for fault boundary identification.[32]

The variance attribute constructs a variance volume by calculating the amplitude variance between each seismic trace in the trace gather and the average seismic trace. Its core lies in characterizing the differences in reflection characteristics between adjacent seismic traces (mainly reflected in amplitude variations), thereby exploring abnormal seismic reflection information caused by anomalous geological bodies or faults, and ultimately achieving the effective identification of faults and fractures.

Figure 8 presents the horizon slice results of 12 variance volumes with distinct dominant frequencies. The comparative analysis demonstrates that within the low-frequency band (5–20 Hz), the 20 Hz variance volume provides the clearest delineation of large-scale fractures. Additionally, within the medium-frequency band (25–40 Hz), the 35 Hz variance volume achieves optimal performance in identifying medium-scale fractures. In the high-frequency band (45–60 Hz), the 50 Hz variance volume shows the best effect in characterizing densely distributed small-scale fractures. Accordingly, these three single-frequency variance data volumes are ultimately selected as the input data for subsequent high-quality frequency-divided ant-tracking.

### 3.5. Attitude-constrained frequency-divided ant-tracking

The identification performance of the ant-tracking

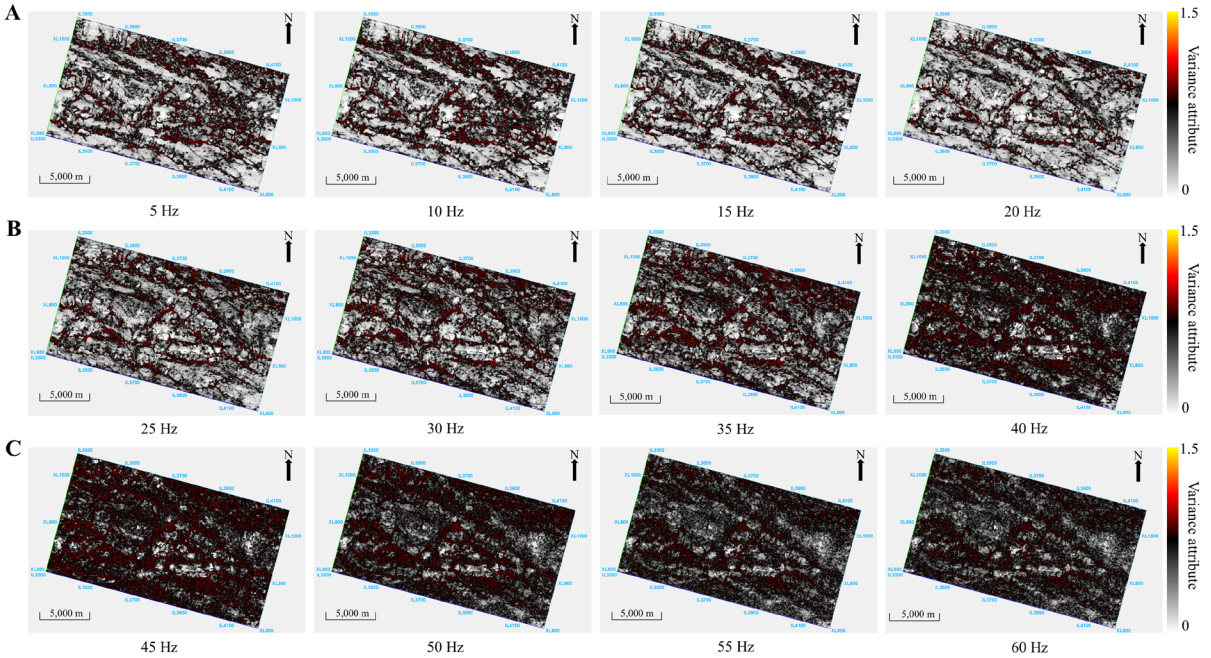



**Figure 8.** Horizon slices of discontinuity detection for frequency-divided seismic data volumes (scale bar = 5,000 m): (A) Low-, (B) medium-, and (C) high-frequency bands.

algorithm is mainly jointly constrained by algorithm parameter settings and regional structural background characteristics. In this study, the active ant-tracking algorithm was adopted. This algorithm focuses on refined search in local areas, features a short step length and low deviation tolerance, and is more suitable for the precise characterization of fracture systems under complex structural backgrounds. After optimizing and selecting the ant-tracking parameter combination, it is still necessary to introduce an attitude constraint mechanism. Without the constraint of regional structural characteristics, ant-tracking will perform omnidirectional, undifferentiated fracture search, but fracture development in the actual study area has significant directionality and is not randomly distributed in all directions, making the tracking results prone to false fractures. The attitude-constrained ant-tracking method is thus a key means to eliminate false fracture information and improve identification reliability.

The specific implementation process is as follows: based on the generated single-frequency seismic attribute data volumes, combined with previous research findings, key attitude parameters such as the strike, dip direction, and dip angle of major faults in the area were statistically analyzed to clarify the spatial distribution patterns of fracture development, and targeted filtering criteria were subsequently formulated. Meanwhile, considering that horizon reflection traces usually exhibit low-dip angle responses, low-dip angle anomalies need to be filtered out. During the actual ant-tracking process, precise control of the search direction was achieved through the stereographic projection diagram in Petrel E&P software platform (2018, Schlumberger Limited, France; Figure 9). The dip angle and azimuth range of fractures to be filtered out are set as the gray area, and fractures within this range will not be included in the tracking results. In contrast, fractures corresponding to the white area are the target objects of ant-tracking and will be recorded.

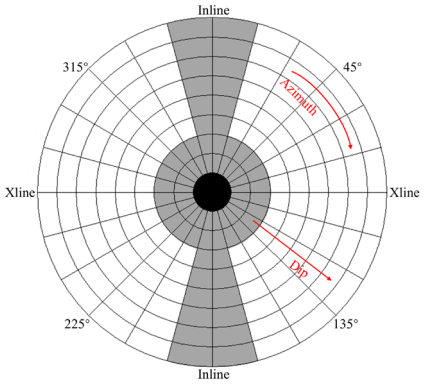



**Figure 9.** Stereonet map generated using attitude-constrained ant-tracking in Petrel E&P software platform

Previous studies have shown that the Raphia buried hill belt extends in a Northwest (NW)–SE direction, with a length of approximately 20 km, a width of about 5 km,

and a structural area of around 100 km$^2$. Controlled by NW-trending major faults, the buried hill belt generally strikes NW–SE, and the mountain tops from South to North exhibit an en-echelon distribution pattern. A NW-trending mountain ridge develops along the line of Raphia SW-2–Raphia S-3–Raphia S-11–Raphia S-9 within the buried hill belt, and the burial depth gradually decreases from NW to SE, with a maximum burial depth of about −500 m, a minimum burial depth of approximately −1,900 m, and a maximum vertical drop of up to 1,400 m, forming multiple fault-block mountains with significant height differences.

Three sets of faults are mainly developed in the Raphia S buried hill belt (Figure 10). The first set is approximately NW–NW-west (NWW) trending faults, which serve as major faults in the area and control the formation and spatial distribution of the buried hill belt. The second set is approximately Eastwest (EW)-trending faults, which function to divide the buried hill belt into blocks and significantly control the development of buried hill fractures. The third set is near NE-trending faults, which cut the NW-trending major faults, complicate the internal structure of the buried hill, and play a key role in promoting the formation and development of buried hill fractures. Detailed characteristics of the main faults in the Raphia buried hill belt are shown in Table 1.

Based on the actual geological conditions of the Raphia S block, the dip angle range of faults participating in the ant-tracking calculation was finally set to 60–90° (as nearly vertical strike-slip faults develop in the area, 90° was selected as the maximum dip angle threshold). In terms of azimuth, NW–NWW, approximately EW, and NE-trending fractures are retained for calculation.

Under attitude constraints, ant-tracking was performed on the single-frequency variance data volumes of 20 Hz, 35 Hz, and 50 Hz, respectively, with the results shown in Figure 11. The 20 Hz low-frequency ant volume (Figure 11A) clearly depicts the planar distribution of large-scale NW–NWW trending faults that control the buried hill framework. The 35 Hz medium-frequency ant

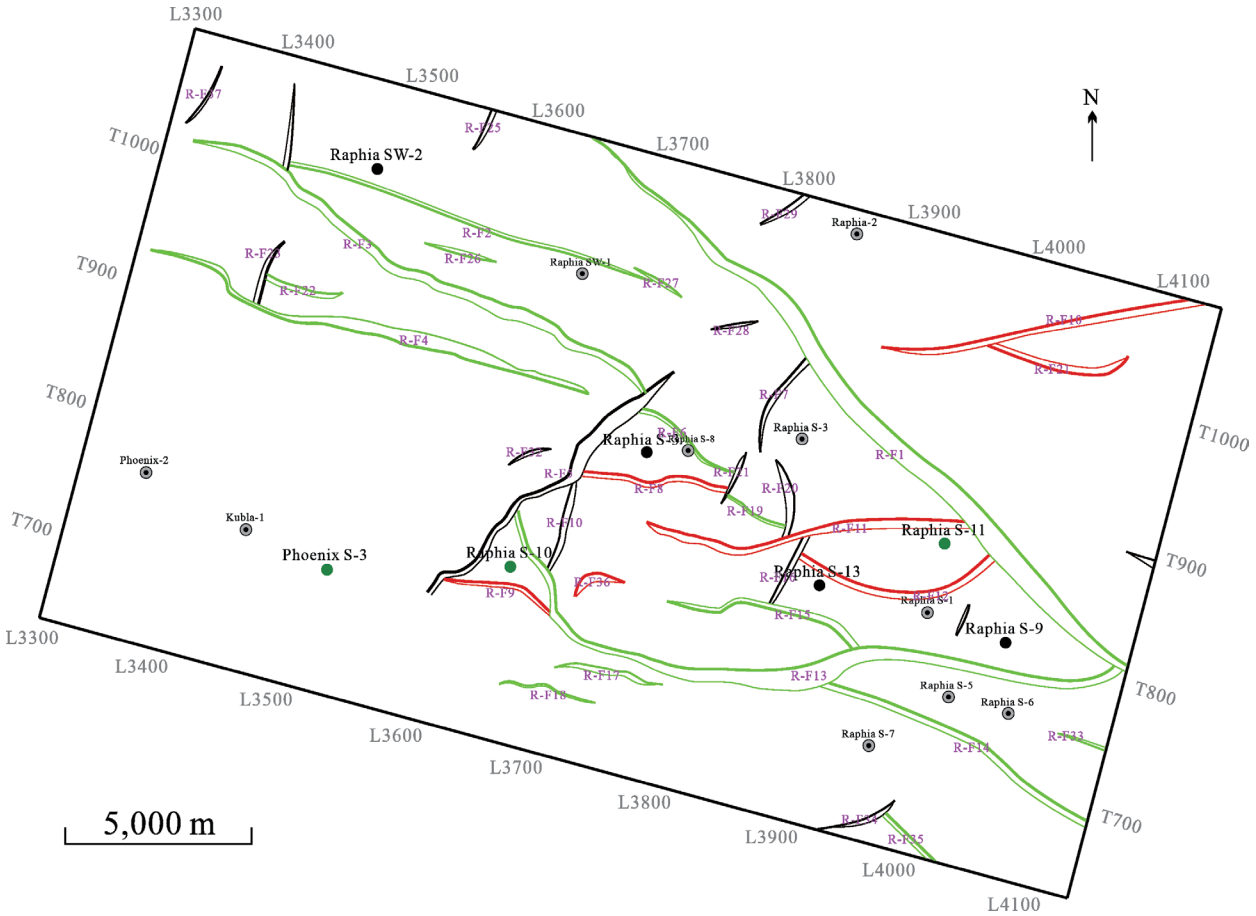



**Figure 10.** Fault distribution map of the Raphia S buried hill belt (scale bar = 5,000 m). The first fault set is highlighted in green, the second set in red, and the third set in black.

**Table 1. Main fault elements of Raphia S block buried hill belt**

| Fault name | Fault type | Strike | Azimuth | Dip angle (°) |
|---|---|---|---|---|
| R-F1 | Normal fault | NW | NE | 65–80 |
| R-F2 | Normal fault | NWW | SSW | 60–80 |
| R-F3 | Normal fault | NW | NE | 70–80 |
| R-F4 | Normal fault | NWW | SSW | 70–85 |
| R-F5 | Normal fault | NE | NW | 70–80 |
| R-F6 | Normal fault | NW | NE | 65–80 |
| R-F7 | Normal fault | NE | NW | 60–80 |
| R-F8 | Normal fault | EW | N | 70–80 |
| R-F9 | Normal fault | EW | S | 65–80 |
| R-F10 | Normal fault | EW | NWW | 65–85 |
| R-F11 | Normal fault | EW | N | 55–70 |
| R-F12 | Normal fault | EW | N | 60–75 |
| R-F13 | Normal fault | EW | N | 65–80 |
| R-F14 | Normal fault | NW | NE | 65–70 |
| R-F15 | Normal fault | NWW | SSW | 60–80 |
| R-F16 | Normal fault | NE | NW | 65–80 |

Abbreviations: EW: Eastwest; N: North; NE: Northeast; NW: Northwest; NWW: Northwest-west; S: South; SSW: South-southwest.

volume (Figure 11B) retains the outline of large faults and reveals more abundant medium-scale fractures, especially enhancing the response to near-EW and NE trending fractures. The 50 Hz high-frequency ant volume (Figure 11C) further identifies a large number of densely distributed small-scale fractures with the richest detailed information. This frequency-divided processing strategy effectively achieves the separation and identification of multi-scale fracture systems from large-scale to small-scale fractures.

## 4. Application results

To systematically verify the high precision, high resolution, and refined characterization capability of the frequency-divided ant-tracking method based on WT-ISD in fracture prediction, three fracture prediction approaches were selected for application and comparative analysis. These approaches include conventional ant-tracking, frequency-divided ant-tracking based on conventional ST, and frequency-divided ant-tracking based on WT-ISD. Comprehensive comparative verification was conducted from both seismic profile and planar dimensions.

Before the verification, RGB attribute fusion was first performed on three sets of single-frequency ant volume data. These data were obtained through frequency division processing via ST and WT-ISD, with each set corresponding to a distinct dominant frequency. RGB fusion is not a simple image superposition—its scientific validity depends on the assignment of reasonable weights. Correlation analysis was conducted between the fault polygons of the Raphia S block and the ant volumes at dominant frequencies of 20 Hz, 35 Hz, and 50 Hz. A higher correlation corresponds to a larger weight. The correlation analysis results indicate that the weight coefficients of the 20 Hz, 35 Hz, and 50 Hz ant volumes are 0.5, 0.35, and 0.15, respectively. Based on this weight scheme, attribute-weighted fusion was completed

for the three sets of single-frequency ant volumes with different dominant frequencies.

In the profile verification stage, three sets of results were superimposed on the original seismic profile for comparative display, including the conventional ant-tracking results (Figure 12A), the attribute-fused frequency-divided ant-tracking results based on conventional ST (Figure 12B), and the attribute-fused frequency-divided ant-tracking results based on WT-ISD (Figure 12C). This display is designed to visually reveal the spatial distribution characteristics of the fault system. As illustrated in Figure 12, the consistency between fractures identified by conventional ant-tracking and seismic reflection features is significantly lower than that of the frequency-divided ant volume results integrated across multiple frequencies. Compared with the attribute-fused ant volumes derived from ST-based frequency division, the attribute-fused ant volumes obtained from WT-ISD–based frequency division exhibit superior performance. They can characterize fractures of all scales more clearly, identify a notably greater number of fractures, and present more complete continuity of fault structures in the profile space.

In the planar characteristic analysis stage, the WT-ISD–based frequency-divided ant-tracking results processed via RGB attribute fusion were visualized. Finally, a comprehensive fracture prediction map of the study area was generated (Figure 13). This fused result effectively integrates the seismic response characteristics of fractures at different scales across all frequency bands. It also highlights the complementary contribution of multi-frequency information in fracture detection. Accordingly, a more comprehensive and refined characterization of the fracture network in the study area is achieved.

To further verify the superiority of the proposed WT-ISD–based frequency-divided ant-tracking method, a systematic comparison was conducted among three sets of results—the conventional ant-tracking results (Figure 14A), the ST-based frequency-divided ant-tracking results (Figure 14B), and the WT-ISD–based frequency-divided ant-tracking results (Figure 14C).

The analysis demonstrates that the prediction outcomes of the WT-ISD–based frequency-divided ant-tracking method are significantly superior to those of the other two methods across multiple evaluation indicators. In terms of the richness of fracture identification, frequency division processing effectively separates the frequency response characteristics of fractures at different scales. This enables the ant volumes to clearly characterize large-, medium-, and small-scale fracture systems. The number of identified fractures increases markedly, and the structural framework

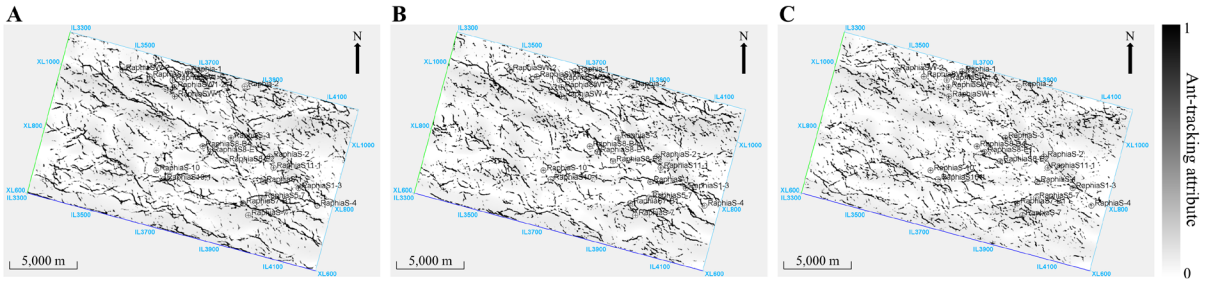



**Figure 11.** Horizon slices of attitude-constrained frequency-divided ant-tracking for the buried hill horizon (scale bar = 5,000 m): (A) 20 Hz ant volume, (B) 35 Hz ant volume, and (C) 50 Hz ant volume

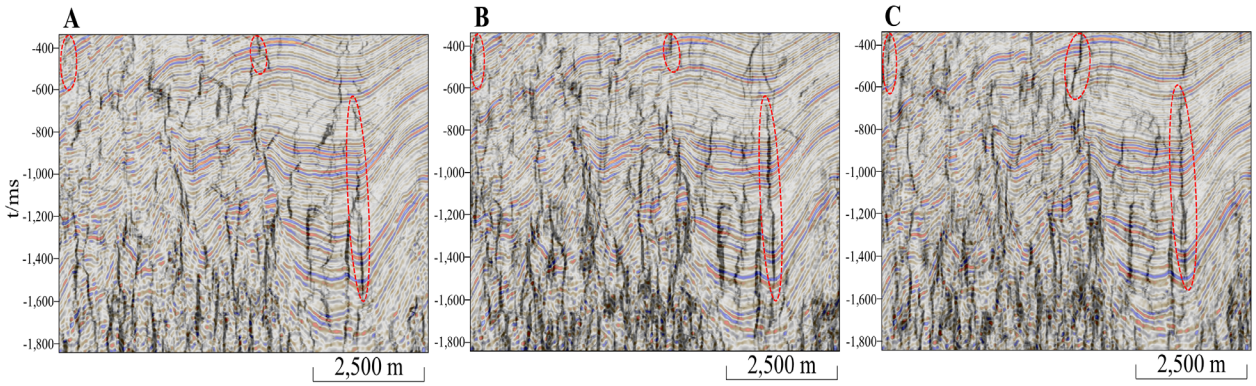



**Figure 12.** Comparative verification of ant-tracking results along the seismic profile (scale bar = 2,500 m). (A) Conventional ant-tracking. (B) Frequency-divided ant-tracking based on conventional S-transform. (C) Frequency-divided ant-tracking based on W-transform inversion spectral decomposition.

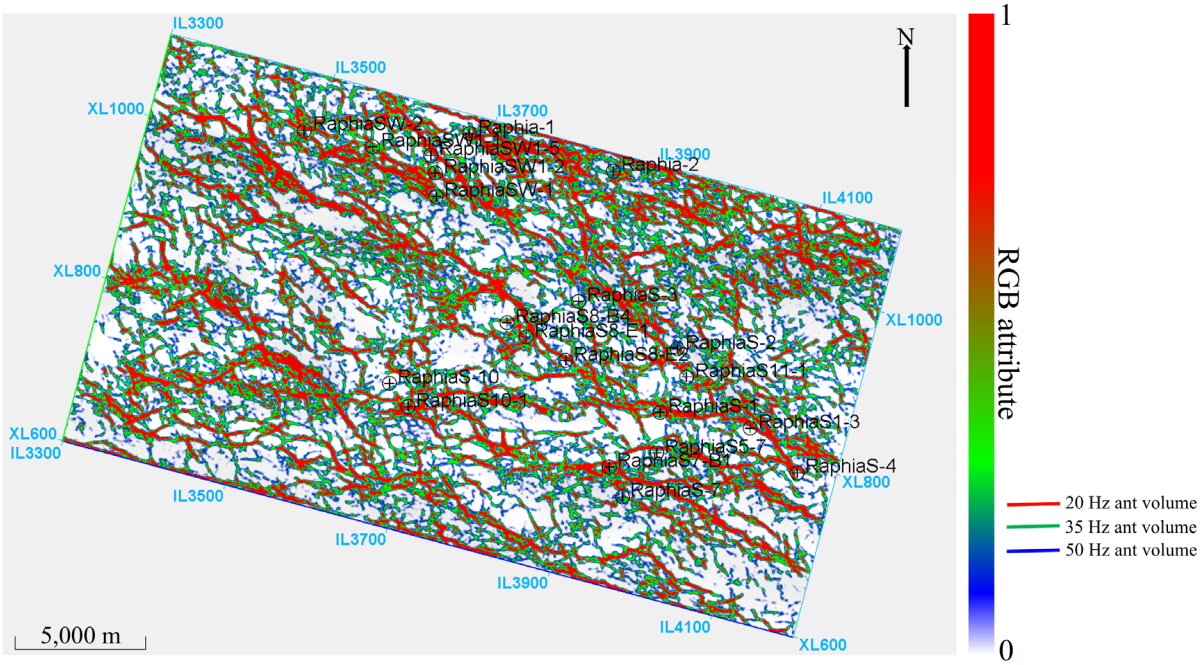



**Figure 13.** Red–green–blue (RGB) attribute fusion results of W-transform inversion spectral decomposition-based frequency-divided ant-tracking (scale bar = 5,000 m)

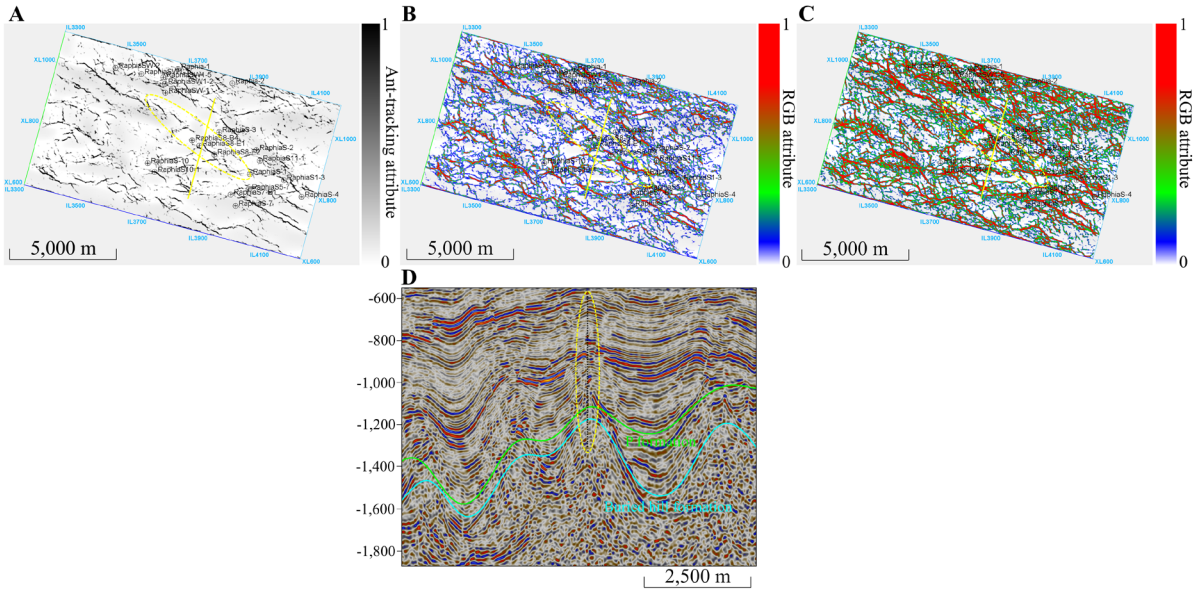



**Figure 14.** Comparison of three ant-tracking fracture prediction methods. (A) Conventional ant-tracking (scale bar = 5,000 m). (B) S-transform–based frequency-divided ant-tracking (scale bar = 5,000 m). (C) W-transform inversion spectral decomposition-based frequency-divided ant-tracking (scale bar = 5,000 m). (D) Seismic profile across the study area showing the P formation and buried hill formation, where the fracture system within the yellow dashed circle corresponds to the fracture system marked by the yellow dashed circle in the planar results of (A)–(C) (scale bar = 2,500 m).

of the fracture system becomes more complete.

In terms of the spatial continuity of faults, the fault distribution presented by the frequency-divided ant volumes exhibits higher coherence and integrity. The continuous identification ability is significantly enhanced, especially when tracking major fault systems with a wide extension range. In terms of consistency with the structural strike, the frequency-divided results are highly consistent

with the fault distribution characteristics of the buried hill belt in the Raphia S block, indicating strong geological conformity. Notably, in the identification of NW–NWW-trending major faults, the frequency-divided ant volumes obtained based on WT-ISD have clear boundaries and regular morphologies. This represents a remarkable improvement over the problems of blurred fault boundaries and insufficient continuity in the other two methods. This improvement effectively enhances the interpretability and reliability of the fracture prediction results.

## 5. Conclusion

In this study, a frequency-divided ant-tracking method based on WT-ISD was systematically developed and validated, successfully overcoming the limitations of conventional ant-tracking, such as sensitivity to data quality and poor continuity of fault identification. Comprehensive tests on actual seismic data from the Raphia S Block in the Bongor Basin yielded the following key conclusions:

(i) The WT-ISD achieves substantial advancements in algorithm design and functional performance for time–frequency analysis. Algorithmically, it innovatively integrates sparse inversion theory with the WT, constructing an overcomplete dictionary matrix and introducing L1-norm sparsity constraints. Solved via the SPGL1 algorithm, this framework transforms traditional linear time–frequency analysis into a model-driven nonlinear inversion process. Functionally, WT-ISD effectively breaks the inherent time–frequency resolution trade-off imposed by the Heisenberg uncertainty principle that constrains conventional linear methods. It delivers time–frequency spectra with superior energy concentration, higher resolution, and enhanced noise resistance, enabling precise characterization of time–frequency responses from multi-scale geological bodies (e.g., fractures, thin layers). Beyond fracture prediction, this advancement provides a robust tool for high-fidelity non-stationary signal decomposition, with broad implications for geophysical signal processing and related interdisciplinary fields.

(ii) The frequency-divided fracture prediction workflow exhibits strong generalizability and excellent multi-scale fracture identification capability. Application results confirm that extracting single-frequency data volumes sensitive to 20 Hz (large-scale fractures), 35 Hz (medium-scale fractures), and 50 Hz (small-scale fractures) enables effective separation and refined characterization of multi-scale fracture systems. The final RGB-fused fracture model outperforms conventional methods in fracture richness, spatial continuity, and consistency with regional structures. This workflow is not limited to the Bongor Basin; its data-driven design allows flexible adjustment of key parameters (e.g., frequency bands, attitude constraints, fusion weights) to adapt to different basin types (e.g., rift basins, foreland basins) and fracture systems (e.g., tectonic, diagenetic, hybrid), providing a practical solution for accurate multi-scale fracture characterization in complex reservoirs.

## Acknowledgments

None.

## Funding

This research was supported by the Open Fund Project (Tenth Batch) (Grant No. K2024-03) of the Key Laboratory of Oil and Gas Resources and Exploration Technology, Ministry of Education, Yangtze University, China.

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Ying Jia, Xilin Qin
*Formal analysis:* Xiuwei Wang, Xilin Qin
*Funding acquisition:* Ying Jia, Xiuwei Wang
*Investigation:* Ying Jia, Jun Wang
*Methodology:* Xilin Qin, Yinglong Li
*Project administration:* Ying Jia, Xiuwei Wang
*Validation:* Yinglong Li
*Visualization:* Xilin Qin, Yinglong Li
*Writing–original draft:* Ying Jia, Yinglong Li
*Writing–review & editing:* Ying Jia, Xilin Qin

## Availability of data

Data are available from the corresponding author upon reasonable request.

## References

1. Liu YW, Liu XW, Han L, Liu ZY. Development status of seismic fracture prediction technology. *World Pet Ind*. 2024;31(3):26-34. [In Chinese].

   doi: 10.20114/j.issn.1006-0030.20240130001

2. Du X, Fan TE, Fan PJ, Ma TF, Zhang XW. Seismic prediction method of multi-scale fracture zone for buried-hill fractured reservoir. *Chin J Geophys*. 2024;67(8):3184-3195. [In Chinese].

   doi: 10.6038/cjg2024R0451

3. Zhu BH. Improved Ant Tracking Crack Detection Algorithm and Its Application. *Offshore Oil*. 2019;39(3):27-32. [In Chinese].

doi: 10.3969/j.issn.1008-2336.2019.03.027

4. Liang ZQ. Poststack seismic prediction techniques for fractures of different scales. *Geophys Prospect Pet*. 2019;58(5):766-772.

   doi: 10.3969/j.issn.1000-1441.2019.05.016

5. Portnoff M. Time-frequency representation of digital signals and systems based on short-time Fourier analysis. *IEEE Trans Acoust Speech Signal Process*. 1980;28(1):55-69.

   doi: 10.1109/tassp.1980.1163359

6. Uesugi F. Novel image processing method inspired by wavelet transform. *Micron*. 2023;168:103442.

   doi: 10.1016/j.micron.2023.103442

7. Stockwell RG, Mansinha L, Lowe RP. Localization of the complex spectrum: the S transform. *IEEE Trans Signal Process*. 1996;44(4):998-1001.

   doi: 10.1109/78.492555

8. Zhao ZC, Rao Y, Wang YH. Structure-adapted Multichannel Matching Pursuit for Seismic Trace Decomposition. *Pure Appl Geophys*. 2023;180(3):851-861.

   doi: 10.1007/s00024-023-03234-w

9. Daubechies I, Lu J, Wu HT. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl Comput Harmon Anal*. 2011;30(2):243-261.

   doi: 10.1016/j.acha.2010.08.002

10. Zhang GL, Duan J, Li Y, He CJ, Du H, Luo F. Adaptive Time-Resampled High-Resolution Synchrosqueezing Transform and Its Application in Seismic Data. *IEEE Trans Geosci Remote Sens*. 2020;58(9):6691-6698.

    doi: 10.1109/TGRS.2020.2978509

11. Partyka G, Gridley J, Lopez J. Interpretational applications of spectral decomposition in reservoir characterization. *Lead Edge*. 1999;18(3):353-360.

    doi: 10.1190/1.1438295

12. Zeng H, John A, G JK. Frequency-dependent Seismic Stratigraphy. In: *Proceedings of the SEG Technical Program Expanded Abstracts 2009.* 79th SEG Annual International Meeting; October 5–30, 2009; Houston, Texas, USA. Society of Exploration Geophysicists; 2009:1097-1101.

    doi: 10.1190/1.3255040

13. Wang YH. The W transform. *Geophysics*. 2021;86(1):31-39.

    doi: 10.1190/geo2020-0316.1

14. Li R, Zhu X, Zhou YZ, Chen H, Chen XP, Hu Y. Generalized W transform and its application in gas-bearing reservoir characterization. *IEEE Geosci Remote Sens Lett*. 2021;19:1-5.

    doi: 10.1109/LGRS.2021.3108836

15. Chen H, Peng L, Chen XP, Hu Y, Li R, Fang YX. A three-parameter W transform and its application to gas reservoir identification. *Geophysics*. 2022;87(5):V521-V532.

    doi: 10.1190/geo2021-0803.1

16. Luo CP, Zong ZY. The synchroextracting algorithm based on W transform and its application in channel characterization. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

    doi: 10.1109/LGRS.2023.3262637

17. Chen XP, Hu Y, Chen H, Wen XT, Guo S. The Normalized W Transform for Seismic Interpretation. *Geophysics*. 2025;90(5):V421-V438.

    doi: 10.1190/geo2024-0678.1

18. Dorigo M, Maniezzo V, Colorni A. Ant system: optimization by a colony of cooperating agents. *IEEE Trans Syst Man Cybern B Cybern*. 1996;26(1):29-41.

    doi: 10.1109/3477.484436

19. Yuan YO, Simons FJ. Multiscale adjoint waveform-difference tomography using wavelets. *Geophysics*. 2014;79(3):WA79-WA95.

    doi: 10.1190/geo2013-0383.1

20. Wen H, Wang Z, Ding LM, *et al*. Fracture detection method of messy body attribute based on multichannel synchro squeezing generalized S-transform. *Prog Geophys*. 2025;40(3):992-1003. [In Chinese].

    doi: 10.6038/pg2025HH0590

21. Liu Q, Wang YX, Xu YG. Synchrosqueezing extracting transform and its application in bearing fault diagnosis under non-stationary conditions. *Measurement*. 2021;173:108569.

    doi: 10.1016/j.measurement.2020.108569

22. Xie QH, Zhao CD, Rui ZF, Guan SJ, Zheng WT, Fan HC. An improved ant-tracking workflow based on divided-frequency data for fracture detection. *J Geophys Eng*. 2022;19(5):1149-1162.

    doi: 10.1093/jge/gxac075

23. Yu C, Wang SY, Cheng LF. Application of high-resolution processing in seismic data based on an improved synchrosqueezing transform. *Front Earth Sci*. 2022;10:956817.

    doi: 10.3389/feart.2022.956817

24. Zhang HX, Wu XF, He BS, Guo M. Time-synchroextracting of generalized S-transform and its application in fault identification. *J Appl Geophys*. 2023;216:105147.

    doi: 10.1016/j.jappgeo.2023.105147

25. Wang Q, Li YX, Chen SJ, Tang B. Matching Demodulation Synchrosqueezing S Transform and its Application in Seismic Time-Frequency Analysis. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

    doi: 10.1109/LGRS.2020.3047892

26. Tian L. Seismic spectral decomposition using short-time

fractional Fourier transform spectrograms. *J Appl Geophys*. 2021;192:104400.

doi: 10.1016/j.jappgeo.2021.104400

27. Chen YP, Peng ZM, Cheng ZY, Tian L. Seismic signal time-frequency analysis based on multi-directional window using greedy strategy. *J Appl Geophys*. 2017;143:116-128.

doi: 10.1016/j.jappgeo.2017.05.017

28. Qin X, Chen S, Zhang S, Li X, Liu Y. The application of frequency-dependent AVO inversion in tight reservoirs area. In: *Proceedings of the SEG Technical Program Expanded Abstracts 2016*. 86th SEG Annual International Meeting; October 21, 2016; Dallas, Texas, USA. Society of Exploration Geophysicists; 2016:3548-3552.

doi: 10.1190/segam2016-13839860.1

29. Yu ZH, Xiao KY, Zhang GL, Xiao GJ, Du YB. Analysis on inverted structure characteristics and its forming mechanism in the Bongor Basin, Chad. *China Pet Explor*. 2018;23(3):90-98. [In Chinese].

doi: 10.3969/j.issn.1672-7703.2018.03.011

30. Li SJ, Wang TY, Gao JH, Liu BY, Gui JY, Wang HQ. lp norm inverse spectral decomposition and its multi-sparsity fusion interpretation. *Appl Geophys*. 2021;18(4):569-578.

doi: 10.1007/s11770-021-0919-x

31. Zheng H, Cai JX, Wang JB. Gaussian beam tomography with structure-filtering and its applications. *Geophys Geochem Explor*. 2020;44(2):372-380. [In Chinese].

doi: 10.11720/wtyh.2020.1300

32. Xie QH, Jiang LW, Zhao CD, Wang ZD, Tang XH, Luo YF. Application study of improving the precision of the ant-tracking-based fracture prediction technique. *Geophys Geochem Explor*. 2021;45(5):1295-1302. [In Chinese].

doi: 10.11720/wtyh.2021.1208

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Interpretable intelligent gas-bearing reservoir prediction using time–frequency analysis and manifold-regularized semi-supervised GAN

**Shuying Ma[1], Junxing Cao[2,3]*, Rong Wang[1]*, Xudong Jiang[3], Jun Wang[2,3], Lingsen Zhao[3], Hong Li[3], and Xin Tang[3]**

[1]School of Network and Communication Engineering, Chengdu Technological University, Chengdu, Sichuan, China

[2]State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu, Sichuan, China

[3]College of Geophysics, Chengdu University of Technology, Chengdu, Sichuan, China

## Abstract

As hydrocarbon exploration advances toward deep and complex reservoirs, the identification accuracy of traditional time–frequency (TF) analysis is constrained by strongly heterogeneous geological conditions. Concurrently, while deep learning has shown great potential, mainstream supervised models commonly face the dual challenges of scarce labeled samples and the lack of interpretability in their black-box decision-making processes. To address these challenges, this study proposes an innovative, intelligent prediction framework integrating high-precision TF analysis, manifold-regularized semi-supervised generative adversarial networks, and SHapley Additive exPlanations (SHAP) for interpretability. First, the Fourier-based synchrosqueezing transform was employed to extract two-dimensional TF features with superior energy concentration, effectively overcoming the resolution limits imposed by the Heisenberg uncertainty principle. Subsequently, the manifold-regularized semi-supervised generative adversarial network was developed. By incorporating manifold regularization constraints, the discriminator captures the intrinsic topological structure of large-scale unlabeled samples, effectively leveraging data geometry to significantly enhance generalization capability under sparse-label conditions. Finally, the SHAP method was utilized to conduct a post hoc interpretation. Experimental results on the Marmousi II model demonstrate a remarkable testing accuracy of 98.4%. In a real-world application to deep marine reservoirs in the Sichuan Basin, the framework achieved an 85.0% testing accuracy using only 5% labeled samples. Compared to baseline models, the semi-supervised strategy and manifold regularization contributed accuracy gains of 18.8% and 5.0%, respectively. SHAP analysis further confirms the model's adaptive capability to extract geophysical features, enabling it to deconstruct the tuning-effect patterns in synthetic data and the low-frequency enhancement/high-frequency attenuation patterns in real data, respectively. This validation of geophysical consistency provides a theoretical foundation for the application of artificial intelligence in complex hydrocarbon exploration.

***Keywords:*** Gas-bearing prediction; Semi-supervised learning; Synchrosqueezing transform; Generative adversarial network; SHapley Additive exPlanations

***Corresponding authors:**
Junxing Cao
(caojx@cdut.edu.cn)
Rong Wang
(wrong1@cdtu.edu.cn)

**Citation:** Ma S, Cao J, Wang R, *et al*. Interpretable intelligent gas-bearing reservoir prediction using time–frequency analysis and manifold-regularized semi-supervised GAN. *J Seismic Explor*. 2026;35(2):025450102. doi: 10.36922/JSE025450102

**Received:** November 5, 2025

**Revised:** January 9, 2026

**Accepted:** February 10, 2026

**Published online:** April 24, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

As global hydrocarbon exploration shifts toward deep, ultra-deep, and complex hidden reservoirs, seismic signal identification and fluid prediction face unprecedented challenges. When seismic waves propagate through fluid-saturated porous media, they undergo significant energy absorption, scattering, and attenuation due to the combined effects of fluid viscosity and reservoir heterogeneity. These phenomena typically manifest as localized energy anomalies and dispersion signatures in the time–frequency (TF) domain, forming the theoretical foundation for geophysical gas-bearing identification.

Traditional identification methods primarily rely on amplitude anomalies in the time domain (e.g., bright spots[1], dim spots[2], and flat spots[3]) or characteristic changes in the frequency domain (e.g., low-frequency shadows[4], fluid mobility indicators[5], and high-frequency attenuation gradients[6]). While these qualitative methods have achieved some success in simple geological scenarios such as shallow and thick reservoirs, they struggle to support high-reliability quantitative evaluation for deep and thin interbedded targets due to low signal-to-noise ratios (SNR), tuning effects, and strong multi-solution ambiguity.

In recent years, the data-driven paradigm centered on deep learning has demonstrated significant application potential in the intelligent detection of gas-bearing reservoirs, owing to its exceptional capability to uncover non-linear patterns. Numerous studies have attempted to incorporate architectures such as long short-term memory[7,8], deep neural networks[9,10], and convolutional neural networks[11], combined with multiple seismic attributes[12,13], achieving positive progress in specific geological scenarios.

Nevertheless, existing deep learning approaches face substantial challenges in practical production applications: on the one hand, certain models exhibit excessive reliance on long sequences or pre-stack data, resulting in substantial computational overhead; on the other, the manual pre-selection of seismic attributes is prone to introducing subjective bias. A more fundamental constraint lies in the fact that existing models predominantly follow a supervised learning paradigm, whose performance is highly dependent on large-scale, high-confidence labeled samples. In actual oil and gas exploration, constrained by high drilling and coring costs, confirmed gas-bearing well-point labels are extremely scarce and sparsely distributed. This severely limits the generalization capability and engineering applicability of supervised learning models.

To overcome the label scarcity bottleneck, semi-supervised learning (SSL)[14] has emerged as a research focus in geophysical intelligent recognition. This approach leverages both limited labeled data and vast unlabeled datasets for training. Among these, semi-supervised models based on generative adversarial networks (GANs)[15,16] demonstrate unique advantages in feature representation learning and data distribution modeling, showing initial promise in seismic phase identification[17], reservoir prediction[18], and lithology prediction.[19]

Nevertheless, achieving efficient and robust detection of gas-bearing reservoirs requires overcoming three major bottlenecks:

(i) Effective characterization of sensitive features: Extracting TF representations from non-stationary raw signals that are most responsive to gas content and most information-rich forms the foundation for determining model learning efficiency.
(ii) Robustness and generalizability of model training: When confronted with local perturbations and strong background noise, enhancing the model's smoothness and stability within the manifold space is paramount.
(iii) Interpretability of decision mechanisms. Oil and gas exploration constitutes a high-risk decision-making domain. If the decision-making mechanisms of a model (i.e., its black-box characteristics) cannot be understood and trusted, the practical value of its predictive results will be significantly diminished.

By addressing the aforementioned challenges, this study proposes a novel intelligent detection framework for gas-bearing reservoirs that integrates high-precision Fourier-based synchrosqueezing transform (FSST)[20], a manifold-regularized semi-supervised GAN (MR-SSGAN), and SHapley Additive exPlanations (SHAP).[21] The core innovations include:

(i) High-precision TF feature extraction strategy: The FSST is introduced, which uses an instantaneous frequency rearrangement method to squeeze distributed energy back onto actual frequency ridges, significantly reducing the TF blurring inherent in conventional analysis.[22,23] This produces two-dimensional (2D) highly concentrated TF features for deep learning.
(ii) Construction of the MR-SSGAN model: The MR-SSGAN network architecture is developed. This model not only uses GANs to create high-quality pseudo-samples for training set augmentation, but it also incorporates a novel manifold regularization[24] term based on the Laplace norm into the discriminator loss function.[25,26] This constraint compels the model's decision boundary to traverse low-density regions within the data distribution while maintaining

smoothness on the data manifold. Consequently, it efficiently extracts topological structural information from vast unlabeled seismic datasets under conditions of minimal labeled samples, significantly enhancing the model's robustness in complex geological environments.

(iii) Verification of physical interpretability: The SHAP method from game theory is employed to generate post hoc explanations for the trained MR-SSGAN model. By quantifying the marginal contribution of frequency- and time-domain features to the final prediction, the geophysical logic underlying the model's decisions is elucidated, validating the consistency of the model's predictions with fundamental rock physics principles.

## 2. Methods

The overall workflow proposed in this paper comprises three complementary modules, as illustrated in Figure 1: (i) data preprocessing: synchrosqueezing transform (SST)-based TF representation; (ii) construction and training of the MR-SSGAN; and (iii) decision attribution analysis by SHAP.

### 2.1. Data preprocessing: Synchrosqueezing transform-based time–frequency representation

Seismic signals are non-stationary signals whose frequency components undergo significant temporal variation. The traditional short-time Fourier transform (STFT) maps one-dimensional (1D) time signals $f_t \in L^2(R)$ into a 2D TF domain through a sliding window function:

$$STFT_f(f,t) = \int f(t) w(t-\tau) e^{-i2\pi ft} dt \quad (1)$$

where $t$ is time, $f$ is frequency, and $w(t)$ is the window function. The energy distribution in the STFT is significantly influenced by the window function, exhibiting a blurring effect. According to Heisenberg's uncertainty principle, optimal time resolution and frequency resolution cannot be achieved simultaneously, thereby limiting its ability to provide a detailed characterization of complex reservoir signals.

Although the continuous wavelet transform (CWT) incorporates multiscale analysis, it remains constrained in temporal resolution at low frequencies and frequency resolution at high frequencies, while also being sensitive to the selection of wavelet basis functions.

To address the energy blurring effect, Daubechies *et al.*[20] proposed the SST. As a post-processing reassignment technique based on signal phase derivatives, SST squeezes energy dispersed across the TF plane back onto the true instantaneous frequency ridge. For a signal $f_t \in L^2(R)$, its

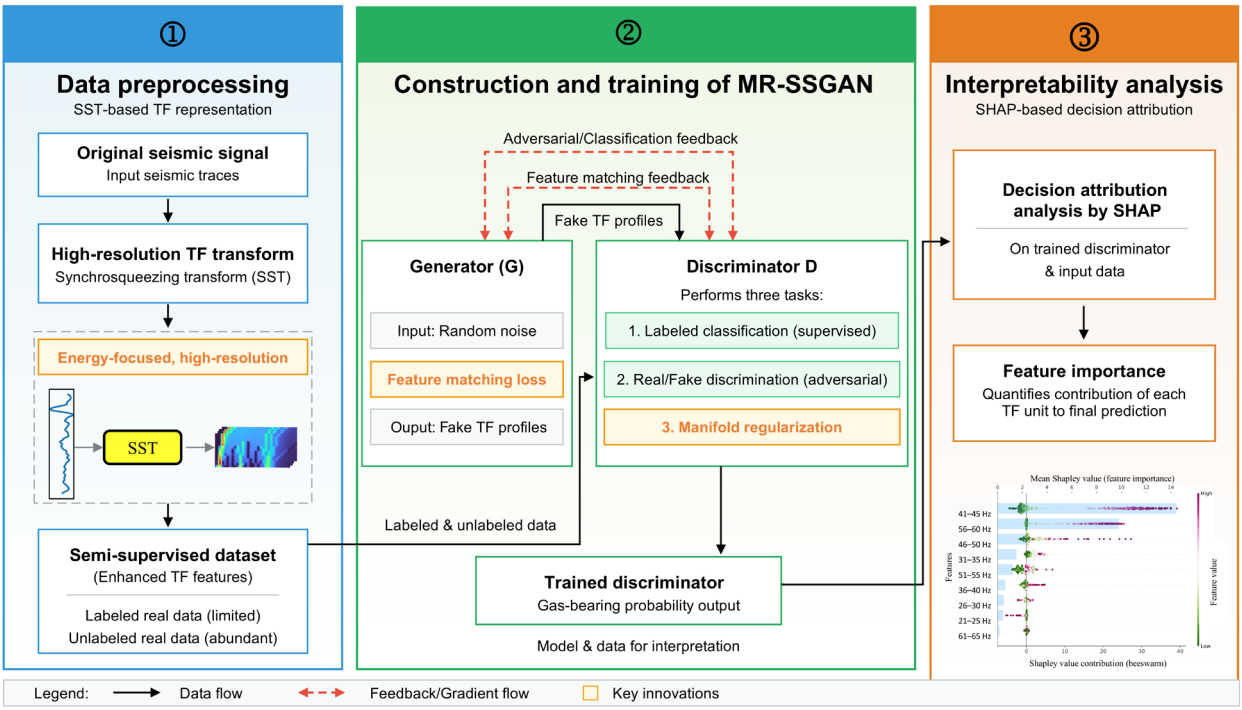



**Figure 1.** Workflow of the intelligent gas-bearing reservoir prediction framework
Abbreviations: MR-SSGAN: Manifold-regularized semi-supervised generative adversarial network; SHAP: SHapley Additive exPlanations; SST: Synchrosqueezing transform; TF: Time–frequency.

instantaneous frequency can be estimated using the partial derivatives of the STFT coefficients:

$$w_f(f,t) = -i\frac{\partial_t \hat{STFT_f}(f,t)}{\hat{STFT_f}(f,t)} \quad (2)$$

where $\hat{STFT_f}(f,t)$ is the complex-valued STFT coefficient. Using this instantaneous frequency estimate, the energy at $(t,f)$ is redistributed to $w_f(f,t)$. For the STFT-based FSST[27], the rearrangement operation is defined as:

$$FSST_f(w_l,t) = \int_{\{f:w_f(f,t)\in F_l\}} \left|STFT_f(f,t)\right| f^{-3/2} df \quad (3)$$

where $F_l$ is the frequency interval and $w_l$ is the center frequency. This operation localizes the originally dispersed energy around the true instantaneous frequency, thereby producing a high-resolution TF spectrum while preserving signal reconstructibility.

### 2.2. Construction and training of the manifold-regularized semi-supervised generative adversarial networks

Standard GANs comprise a generator *G* and a discriminator *D*, which undergo adversarial training through a minimax game. In a semi-supervised setting (SSGAN), the discriminator is transformed into a *K* + 1-class classifier (*K* real classes plus one generated class). For gas-bearing prediction, $K = 2$ (gas-bearing vs. gas-free). *D* must not only determine whether the input data are real or generated, but also predict the physical category (gas-bearing or gas-free) to which genuine samples belong.

Although SSGAN uses unlabeled data, it does not explicitly exploit the topological geometric structure of the data. In contrast, the core innovation of MR-SSGAN is that it fully exploits GANs' potential for modeling natural data manifolds by introducing manifold regularization terms during discriminator *D* training and imposing feature-matching strategies on the generator *D*, resulting in deeper exploitation of the geometric–topological structure of the data.

The loss function $L_D$ for the MR-SSGAN discriminator comprises supervised loss $L_L$, unsupervised loss for genuine unlabeled data $L_U$, unsupervised loss $L_F$ for generated pseudo-data , and the manifold-regularization term $R_{manif}$, as detailed below:

$$L_D = L_L + \omega(L_U + L_F) + \lambda R_{manif} \quad (4)$$

where $\omega = 0.5$ is the weight balancing the contributions of labeled and unlabeled samples, and $\lambda = 10^{-3}$ is the manifold-regularization weight.

For real labeled samples $(x,y)$, the task of the discriminator is to minimize the cross-entropy between the ground-truth labels and the model's predicted distribution. Its loss function $L_L$ is the standard cross-entropy loss:

$$L_L = -E_{x,y\sim p_{data}(x,y)}\left[\log p_D(y \mid x, y < K+1)\right] \quad (5)$$

where $p_D(y \mid x, y < K+1)$ is the probability of the correct label *L* given input *x*.

For real unlabeled data and generated pseudo-data, the corresponding loss terms are unsupervised. $p_D(y = K+1 \mid x)$ provides the probability that input *x* belongs to class *K* + 1 (i.e., pseudo-data). If the input is pseudo-data, the discriminator should identify it as pseudo-data; in this case, the task of the discriminator is to maximize $p_D(y = K+1 \mid x)$. If the input is real unlabeled data, the discriminator should identify it as real data (i.e., the label is not *K* + 1), and its task becomes to minimize $p_D(y = K+1 \mid x)$. Therefore, the corresponding two unsupervised loss functions are defined as follows:

$$L_U = -E_{x\sim p_{data}(x)}\left[\log\left[1 - p_D(y = K+1 \mid x)\right]\right] \quad (6)$$

$$L_F = -E_{x\sim G}\left[\log\left[p_D(y = K+1 \mid x)\right]\right] \quad (7)$$

In the experiment, to simplify the discriminator output while maintaining consistency with the actual number of classes, the property of the Softmax function where subtracting a constant vector from the logits does not change the output probabilities was utilized. The final dimension $l_{K+1}$ of the *K* + 1-dimensional logit vector produced by the discriminator was set to zero, without affecting the final probability outputs.

The manifold hypothesis posits that high-dimensional data are actually distributed on a low-dimensional manifold structure. To explicitly utilize this structure within SSGAN, MR-SSGAN introduces a manifold regularization term $R_{manif}$. According to manifold learning theory, the smoothness of the discriminator can be measured by the integral of the squared gradient over the data manifold (i.e., the Laplacian norm $\|D\|_L^2$). For manifolds parameterized by the generator, samples $z^{(i)}$ may be drawn from the latent representation space *z*. The discriminator's gradient on the manifold can then be estimated via Monte Carlo integration:

$$\|D\|_L^2 \approx \frac{1}{n}\sum_{i=1}^{n}\left\|J_z D\left(G\left(z^{(i)}\right)\right)\right\|_F^2 \quad (8)$$

In **Equation 8**, *n* is the number of samples, which corresponds to the batch size in stochastic gradient descent. *J* represents the Jacobian matrix, used to estimate the

gradient of the discriminator output on the data manifold. $\|\cdot\|_F$ denotes the Frobenius norm. A smaller value for this term indicates smoother discriminator outputs on the manifold.

In practical training, considering the substantial computational overhead associated with directly computing the Jacobian matrix, this study employs a stochastic finite-difference method to approximate this gradient term. If a small perturbation $\varepsilon\bar{\delta}$ is applied to the generator's input $z$, the generated sample should move a small distance along the manifold, at which point the discriminator $D$'s predicted output should remain stable. This yields the final manifold regularization term:

$$R_{manif} = E_{z\sim p_z(z)} d\left(G(z)\right) - D\left(G\left(z+\varepsilon\bar{\delta}\right)\right)_2^2 \tag{9}$$

where $\bar{\delta}$ is the perturbation term, whereas $\bar{\delta} = \frac{\delta}{\|\delta\|}$, $\delta \sim N(0,I)$, and $\varepsilon = 10^{-5}$ represents the perturbation norm.

To address the issues of non-convergence and instability encountered by generators in conventional adversarial training, this study introduces a feature-matching strategy[28] to construct the generator's loss function. The core of this strategy lies in shifting the generator's optimization objective away from solely maximizing the discriminator's misclassification probability. Instead, it seeks to achieve alignment between the distributions of generated and real data within the discriminator's feature space.

Specifically, the expectation of the activation features $h(G[z])$ of the pseudo-samples $G(z)$ generated by the generator at a certain intermediate layer of the discriminator $D$ is designed to match the expectation of the activation features $h(x)$ of the real data $x$ at the same layer. The loss function is defined as follows:

$$L_G = \left\| E_{x\sim p_{data}(x)} h(x) - E_{z\sim p_z(z)} h\left(G(z)\right) \right\|_1 \tag{10}$$

where $h(x)$ represents the activation output of an intermediate layer of the discriminator. In this study, to capture deep and stable semantic representations, the output from the penultimate layer of the discriminator was selected as the feature-matching representation.

The specific network topology and parameter configuration of MR-SSGAN are illustrated in Figure 2. The generator's architecture adopts the classic up-sampling paradigm proposed by Ioffe and Szegedy[29], which has

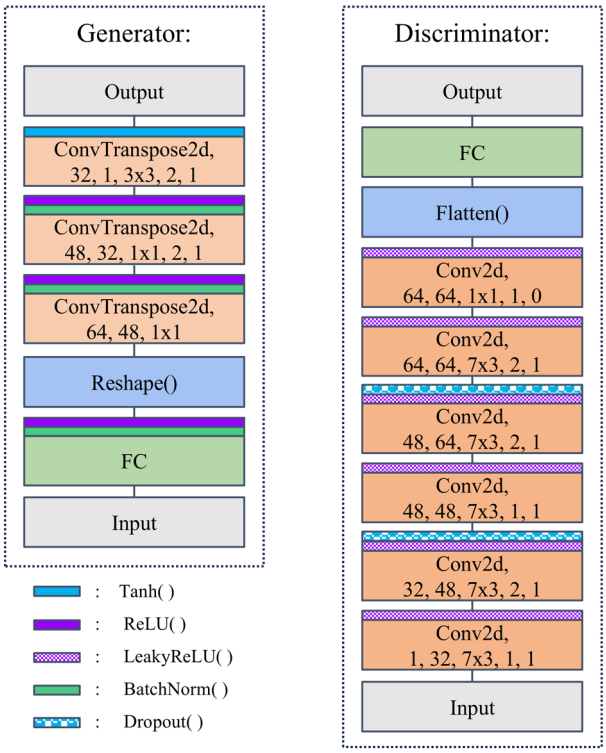



**Figure 2.** Generator and discriminator network architecture of the manifold-regularized semi-supervised generative adversarial network
Abbreviations: BatchNorm: Batch normalization; Conv2d: Convolutional 2D layer; FC: Fully connected layer; ReLU: Rectified linear unit; Tanh: Hyperbolic tangent function.

been adaptively modified to accommodate the input characteristics of 2D TF profiles. It consists of a three-layer transposed convolutional network incorporating batch normalization. In contrast, the discriminator is designed with reference to the robust architecture introduced by Salimans *et al.*[28] for improving GAN training. It is structured as a six-layer 2D convolutional neural network featuring a dropout randomization mechanism, aimed at enhancing the robustness of feature extraction and mitigating overfitting.

By alternately and iteratively optimizing the discriminator loss and the generator loss, the MR-SSGAN framework can efficiently mine potential discriminative information from large-scale unlabeled data while leveraging a small number of critical labeled samples for precise supervision. After training convergence, the discriminator learns the essential characteristic responses of gas-bearing reservoirs in the TF domain, enabling its direct application to gas-bearing prediction of practical 2D seismic profiles.

### 2.3. Decision attribution analysis by SHapley Additive exPlanations

Although the MR-SSGAN model, with its deep non-linear architecture, can capture complex TF features and their interactive effects within seismic signals, its intricate internal parameter mapping results in a decision-making process that exhibits typical black-box properties. To enhance the credibility of model predictions and elucidate its decision-making mechanisms, this study introduces the game-theoretic SHAP attribution analysis.

The core concept of this method is to treat the model's prediction for a specific sample as the result of the combined effects of all input features. By calculating Shapley values, it equitably quantifies the marginal contribution of each feature to the prediction outcome. The computational process follows the classical Shapley formula:

$$\phi_j = \sum_{S \subseteq F \setminus (j)} \left[ \frac{|S|!(M-|S|-1)!}{M!} \times \left[ f\left(S \cup \{j\}\right) - f(S) \right] \right] \tag{11}$$

where $M$ is the total number of features, $F$ is the set of all input features, and $f$ is the discriminator model being explained. $S$ denotes any possible subset derived from the feature set $F$ excluding feature $j$ (i.e., $S \subseteq F \setminus \{j\}$). The difference $f\left(S \cup \{i\}\right) - f(S)$ represents the marginal contribution of adding feature $j$ to the existing subset $S$. By calculating the weighted average of this marginal contribution over all possible subsets $S$, the resulting $\phi_j$ is the Shapley value for feature $j$, which comprehensively measures its contribution to the model's output.

To visually demonstrate the decision-making basis of the MR-SSGAN discriminator for individual gas-bearing prediction, SHAP approximates the original model's decision behavior by constructing a linear local surrogate model $g(x)$. Its additive decomposition expression is as follows:

$$g(x) = \phi_0 + \sum_{j=1}^{p} \phi_j \tag{12}$$

where $x$ is a simplified binary input vector, $P$ represents the total dimensionality of the input TF features, $\phi_0$ denotes the baseline prediction mean for all samples, and $\phi_j$ represents the attribution contribution value corresponding to feature $j$. Through this linear additive decomposition, SHAP can break down complex non-linear decisions into the sum of contributions from each TF unit, thereby clearly identifying the key TF patterns that drive the model's judgment of gas-bearing or gas-free.

## 3. Results

### 3.1. Synthetic data results using the Marmousi II model

This study first validates the algorithm's performance based on the Marmousi II model.[30] Building upon the classical Marmousi model, this iteration incorporates reservoir simulation capabilities, rendering it suitable for assessing the performance of gas-bearing identification algorithms. The core research target comprises a representative gas-bearing sandstone lens located within the common depth point (CDP) 420–550 interval and spanning 1.3–1.4 s, as illustrated in Figure 3A.

To simulate complex field acquisition conditions and assess model robustness, random Gaussian white noise was introduced into the original seismic record, creating a noisy dataset with an SNR of 0 dB. During model training, both the original noise-free data and the 0 dB SNR noisy data were utilized simultaneously to enhance MR-SSGAN's robustness and adaptability to noise.

#### *3.1.1. Time–frequency concentration comparison*

For the CDP 475 single-channel seismic signal, this study compared the feature characterization accuracy of four TF transformation methods (Figure 4). As the performance of CWT is highly dependent on the mother wavelet, the generalized Morse wavelet was selected. Its flexibly adjustable TF window shape enables effective localized characterization. As shown in Figure 4A, although the CWT TF spectrum possesses extremely high temporal resolution, its energy exhibits significant dispersion along the frequency axis.

For the STFT, to overcome the severe spectral leakage

inherent in traditional window functions (e.g., the Hanning window) during detailed analysis, this study employed the discrete phase shell sequence as the window function. This sequence exhibits extremely low sidelobes and rapid sidelobe attenuation characteristics.

Additionally, the findings demonstrated that the frequency resolution of the discrete phase shell sequence-based STFT significantly outperforms that of the CWT (Figure 4B). However, constrained by the Heisenberg uncertainty principle, its temporal resolution remains limited. To further enhance the energy concentration and resolution of the TF representation, the SST was introduced to rearrange the energy in the aforementioned results. Figure 4C, D confirm that both wavelet SST (WSST) and FSST substantially enhance TF concentration: WSST achieves frequency ridge convergence while preserving high temporal resolution, whereas FSST significantly strengthens energy concentration in the temporal dimension while maintaining high frequency resolution.

Further examination of the characterization performance of the aforementioned methods across the profile data was conducted using the 45 Hz frequency-

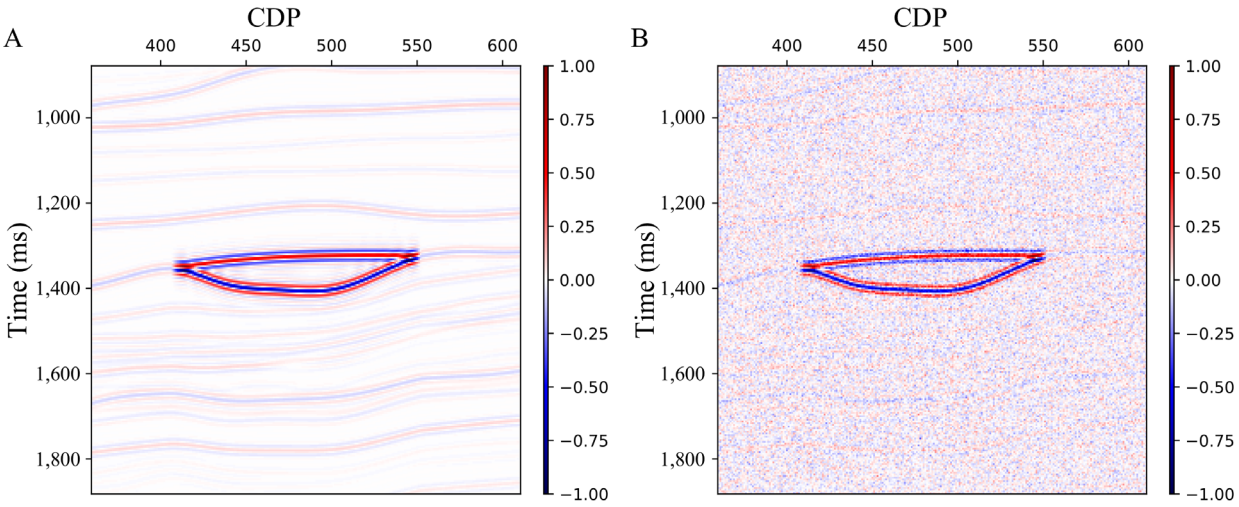



**Figure 3.** Partial data from the Marmousi II model. (A) Clean profile. (B) Noisy profile.
Abbreviations: CDP: Common depth point; SNR: Signal-to-noise ratio.

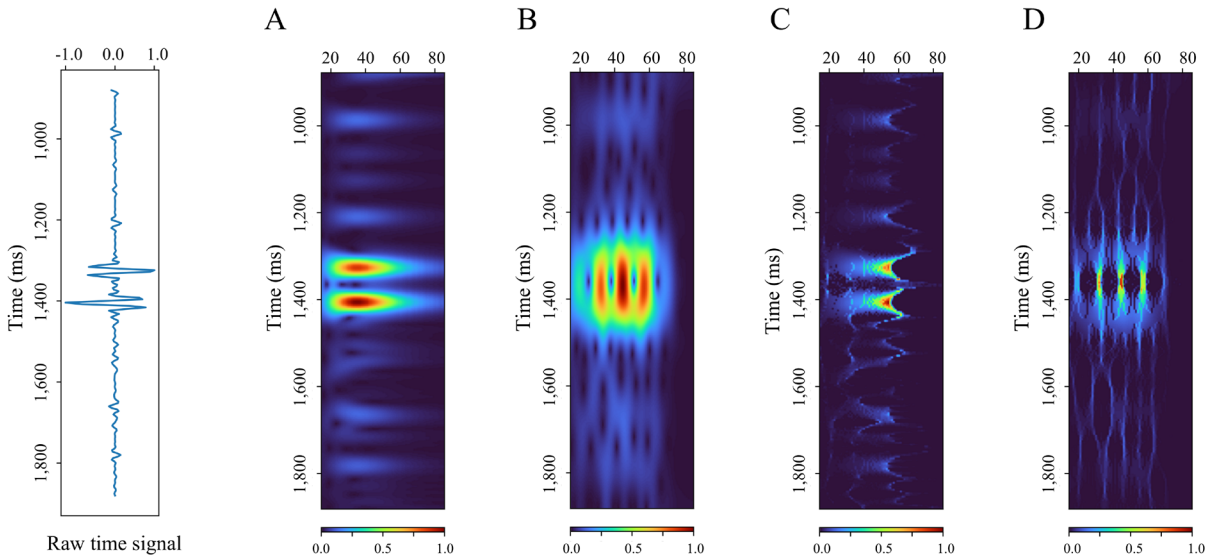



**Figure 4.** The original time-domain signal of the common depth point-475 and its corresponding time–frequency profiles: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

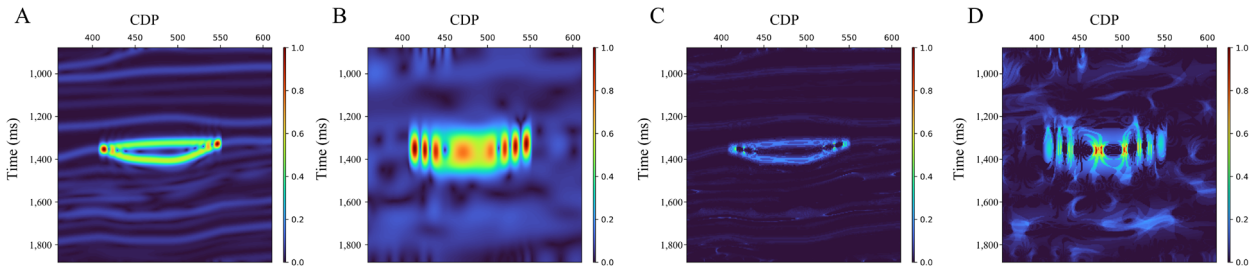



**Figure 5.** 45 Hz frequency slices: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform
Abbreviation: CDP: Common depth point.

separated profile (Figure 5). Results indicate that, while the CWT (Figure 5A) accurately delineates the reservoir top and bottom interfaces, its ability to reveal hydrocarbon enrichment within the reservoir is limited. The STFT (Figure 5B) identifies energy anomaly zones extending far beyond actual reservoir boundaries. Moreover, due to temporal domain blurring effects, the predicted results exhibit pronounced lateral discontinuities and spatial dispersion.

Following synchrosqueezing processing, energy focusing was substantially enhanced, yet identification limitations persisted: WSST (Figure 5C) tended to amplify reservoir boundaries, while FSST (Figure 5D) only displayed high-energy responses within the central reservoir zone. This series of qualitative comparisons demonstrates that, although advanced TF analysis techniques can significantly enhance the physical concentration of seismic characteristics, manual interpretation relying solely on traditional spectral decomposition attributes still suffers from interpretational ambiguity and uncertainty, making it difficult to achieve precise delineation of gas boundaries and intensity.

### *3.1.2. Ablation and robustness analysis*

Traditional gas-bearing prediction based on spectral decomposition relies heavily on manual qualitative interpretation of TF slices. This process is not only highly subjective and inefficient but also struggles to capture subtle, non-linear, sensitive features within seismic signals. In contrast, deep learning frameworks can automatically extract higher-order representations from complex seismic data through multi-layer non-linear mappings. Addressing the challenges of sparse labeling and vast unlabeled datasets in practical geophysical exploration, this study developed an MR-SSGAN model. The model's core advantage lies in leveraging the adversarial interaction between the discriminator and the generator to simultaneously extract prior distribution features from unlabeled data, while strengthening the model's generalization boundaries in sparse labeling scenarios through manifold constraints.

To achieve synchronous representation of TF features, a 2D sliding window sampling strategy $\left(f_{lgh}, t_{lgh}\right)$ was designed to construct 2D TF slice samples. Along the time dimension ($t_{lgh}$), adjacent sampling points before and after the target point are captured, balancing localized temporal information with the integrity of contextual feature learning and preventing degradation in localization accuracy. In the frequency dimension ($f_{lgh}$), the window range was determined based on the instantaneous spectral energy distribution. To capture core signal features while filtering out interference, the frequency interval corresponding to the cumulative energy between the 5th and 95th percentiles at each instantaneous frequency was statistically determined. Through this approach, each sample point was mapped onto a 2D slice sample carrying adjacent spatiotemporal information. Within the experimental configuration of the Marmousi II model (sampling rate of 250 Hz), the extracted TF slice dimensions were set to 45 × 5. The frequency window spans 21–65 Hz ($f_{lgh} = 45$), while the time window covers 20 ms ($t_{lgh} = 5$). This parameter selection is based on statistical analysis of the reservoir response energy accumulation distribution, ensuring coverage of over 90% of the target layer's reflected energy while physically suppressing interference from low-frequency background noise and spurious high-frequency components.

The semi-supervised dataset comprises labeled training data, unlabeled training data, a validation set, and a test set. Based on prior geological information such as logging constraints, gas-bearing and gas-free slices were extracted from well-side traces to form the labeled training set $D_L = \left\{\left(x_i; y_i\right)\right\}_{i=1}^{N_L}$ (sample size $L_{synt}$), while samples distant from the well locations were selected to constitute the unlabeled training set $D_U = \left\{\left(x_j\right)\right\}_{j=1}^{N_U}$ (sample size $U$). The

labels (gas-bearing or gas-free) for 2D TF slices in the labeled samples are determined by the actual gas content at their center points.

Within the Marmousi II synthetic model, the elastic parameters of gas-bearing sandstone lenses are explicitly defined. Gas-bearing samples are defined as sampling points with a $V_p / V_s$ ratio below 1.5 that are located within known sandstone lens boundaries; gas-free samples are defined as sampling points with a $V_p / V_s$ ratio exceeding 1.5 and located outside known sandstone lens boundaries. To ensure model generalization capability and category balance, the ratio of gas-bearing to gas-free samples, as well as raw noise-free to noisy (SNR = 0 dB) samples, was maintained at 1:1. In the experiments, the baseline values for $L_{synt}$, $U_{real}$, and $T_{synt}$ were set to 40, 2,400, and 400, respectively.

To systematically evaluate the impact of labeled sample quantity on model performance, four traces with CDPs of 450, 475, 500, and 510 were selected as training wells. Samples corresponding to quantities of $1\times L_{synt}$, $2\times L_{synt}$, $3\times L_{synt}$, and $4\times L_{synt}$ were sequentially extracted from these wells to form the labeled training set (denoted as $1\times L_{synt}$ to $4\times L_{synt}$), while the unlabeled training and test sets retained their original sample sizes. The labeled and unlabeled training sets collectively constituted the model's training data. Additionally, approximately 10% of the total training samples were selected as the validation set during training, with the discriminator loss monitored in real time using an early stopping strategy.

The model was trained using a learning rate of $5 \times 10^{-4}$ with a batch size of 10, combined with an early stopping strategy (patience = 5) to ensure efficient convergence of the discriminator. To mitigate random fluctuations, all experiments were repeated using three random seeds, and the results were averaged (see Table 1).

Evaluation results demonstrate that, compared to purely supervised GANs, SSGAN effectively enhances the extraction of gas-bearing features by incorporating unlabeled samples into the learning process. Particularly when labeled samples are extremely scarce ($1\times L_{synt}$, i.e., only 40 labeled samples), SSGAN achieves accuracy and F1 scores that are 10.0% and 0.071 higher than those of SGAN, respectively. The performance of MR-SSGAN, which incorporates manifold-regularization constraints, was further improved, achieving additional gains of 4.7% in accuracy and 0.040 in F1 score over SSGAN.

Ablation experiments confirmed that manifold regularization significantly enhances the training stability and generalization of the discriminator under minimal sample sizes by reinforcing the model's predictive smoothness within data manifold neighborhoods. When the labeled data volume increased to $4\times L_{synt}$, MR-SSGAN achieved its optimal performance. Noise resistance analysis conducted under this configuration (Figure 6) demonstrates that the framework maintained high alignment with the ground truth model even under strong background noise at SNR = −5 dB.

To investigate the impact of feature representation quality on classification performance, this study systematically compared the performance differences of four TF transforms—CWT, STFT, WSST, and FSST—as model inputs under fixed data volume ($4\times L_{synt}$-labeled data) (Table 2). The findings revealed that, in the vast majority of cases, the MR-SSGAN model based on FSST input delivers optimal performance, with its predicted profiles (Figure 7) exhibiting the best alignment with the true gas reservoir morphology. This advantage stems primarily from FSST's superior frequency resolution and its more focused energy distribution within the TF domain. These characteristics facilitate the network's learning of clearer feature representations, thereby generating higher-resolution predictive outputs.

In contrast, while STFT also possesses good frequency resolution and its network can partially compensate for temporal resolution limitations through short-time window feature learning, its overall performance was inferior to FSST. CWT and WSST, however, fail to adequately reveal internal details due to energy concentration near reservoir edges in their TF representations (Figure 5), resulting in trained models struggling to accurately identify central reservoir zones. The ablation experiment demonstrates the critical importance of the deep coupling between high-resolution TF characterization and manifold-regularized SSL for enhancing prediction accuracy in gas-bearing reservoirs.

#### *3.1.3. SHapley Additive exPlanations analysis on synthetic data*

To gain deeper insight into how the trained MR-SSGAN model (based on $4\times L_{synt}$-labeled FSST data) makes gas content decisions from 2D TF inputs, this study employed SHAP analysis. SHAP values quantify the contribution of each specific frequency component and time sampling point within the TF data to the prediction outcome for an individual sample. Figure 8 illustrates the feature dependency patterns across both frequency and time dimensions when the model detects gas-bearing states on the Marmousi II model synthetic data, presented through feature importance bar charts and scatter plots.

Analysis revealed that the model assigns the highest

**Table 1. Performance comparison of different models using Fourier-based synchrosqueezing transform data on the Marmousi II model**

| Network architecture | Number of labeled samples | Performance | | Gain vs. SGAN | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| SGAN | $1 \times L_{synt}$ | 82.2 | 0.858 | - | - |
| | $2 \times L_{synt}$ | 86.3 | 0.892 | - | - |
| | $3 \times L_{synt}$ | 91.6 | 0.928 | - | - |
| | $4 \times L_{synt}$ | 95.0 | 0.955 | - | - |
| SSGAN | $1 \times L_{synt}$ | 92.2 | 0.929 | 10.0 | 0.071 |
| | $2 \times L_{synt}$ | 94.1 | 0.947 | 7.8 | 0.055 |
| | $3 \times L_{synt}$ | 96.3 | 0.963 | 4.7 | 0.035 |
| | $4 \times L_{synt}$ | 97.2 | 0.972 | 2.2 | 0.017 |
| MR-SSGAN | $1 \times L_{synt}$ | 96.9 | 0.969 | 14.7 | 0.111 |
| | $2 \times L_{synt}$ | 95.9 | 0.962 | 9.7 | 0.069 |
| | $3 \times L_{synt}$ | 97.5 | 0.976 | 5.9 | 0.048 |
| | $4 \times L_{synt}$ | 98.4 | 0.985 | 3.4 | 0.030 |

Abbreviations: MR: Manifold-regularized; SGAN: Supervised generative adversarial network; SSGAN: Semi-supervised generative adversarial network.

**Table 2. Performance comparison of different time–frequency data using MR-SSGAN on the Marmousi II model**

| Input | Number of labeled samples | Performance | | Loss vs. FSST | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| CWT | $1 \times L_{synt}$ | 95.9 | 0.960 | 0.9 | 0.009 |
| | $2 \times L_{synt}$ | 96.9 | 0.971 | −0.9 | −0.009 |
| | $3 \times L_{synt}$ | 95.9 | 0.962 | 1.6 | 0.014 |
| | $4 \times L_{synt}$ | 96.6 | 0.968 | 1.9 | 0.017 |
| STFT | $1 \times L_{synt}$ | 73.1 | 0.803 | 23.8 | 0.165 |
| | $2 \times L_{synt}$ | 96.3 | 0.964 | −0.3 | −0.003 |
| | $3 \times L_{synt}$ | 97.2 | 0.973 | 0.3 | 0.003 |
| | $4 \times L_{synt}$ | 97.2 | 0.973 | 1.3 | 0.012 |
| WSST | $1 \times L_{synt}$ | 80.6 | 0.817 | 16.3 | 0.151 |
| | $2 \times L_{synt}$ | 83.1 | 0.831 | 12.8 | 0.130 |
| | $3 \times L_{synt}$ | 84.1 | 0.841 | 13.4 | 0.135 |
| | $4 \times L_{synt}$ | 85.3 | 0.855 | 13.1 | 0.129 |
| FSST | $1 \times L_{synt}$ | 96.9 | 0.969 | - | - |
| | $2 \times L_{synt}$ | 95.9 | 0.962 | - | - |
| | $3 \times L_{synt}$ | 97.5 | 0.976 | - | - |
| | $4 \times L_{synt}$ | 98.4 | 0.985 | - | - |

Abbreviations: CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

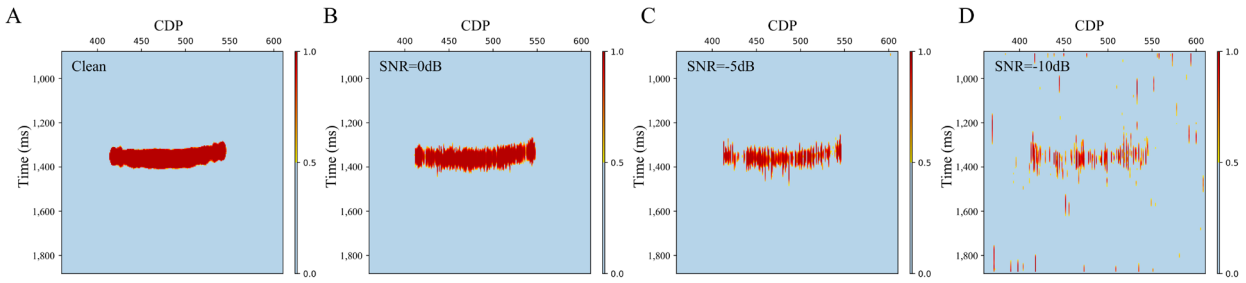



**Figure 6.** Prediction probability of gas-bearing reservoirs by the Fourier-based synchrosqueezing transform manifold-regularized semi-supervised generative adversarial network under varying SNR conditions: (A) clean, (B) 0 dB, (C) −5 dB, and (D) −10 dB
Abbreviations: CDP: Common depth point; SNR: Signal-to-noise ratio.

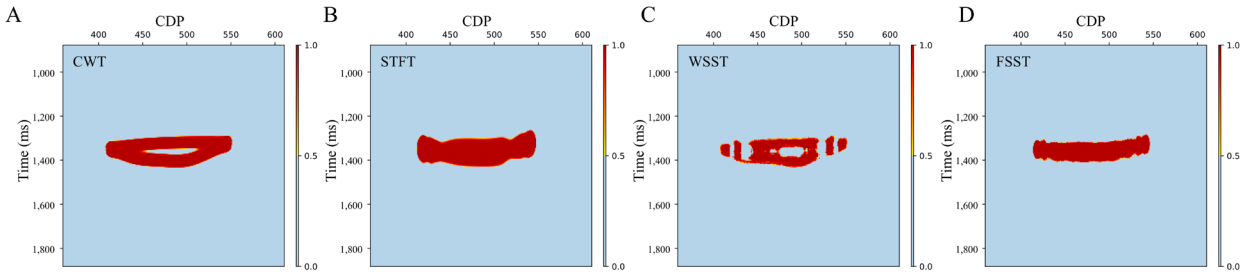



**Figure 7.** 45 Hz frequency slices of synthetic data: (A) CWT, (B) STFT, (C) WSST, and (D) FSST
Abbreviations: CDP: Common depth point; CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

decision weights to the mid-to-high frequency bands (e.g., 41–45 Hz, 46–50 Hz, 51–55 Hz, and 56–60 Hz). When energy in these bands is elevated, SHAP values predominantly cluster on the positive half-axis, strongly favoring gas-bearing predictions; conversely, they tend toward the negative half-axis, indicating gas-free. Notably, the Marmousi II model employs the Ormsby source wavelet (5–80 Hz), with 41–60 Hz falling precisely within its primary frequency response range. The pattern captured by the model, where enhanced energy indicates gas presence, likely stems from relative energy amplification in specific frequency bands within gas-bearing sandstone lenses. This amplification arises from thin-layer tuning effects or constructive wavefield interference. These results demonstrate that MR-SSGAN is not constrained by a single linear attenuation assumption but possesses the capability to uncover complex, non-intuitive seismic physical characteristics.

In the temporal dimension, analysis focused on the current time $T_0$ and five surrounding sampling points ($T_0 - 2$, $T_0 - 1$, $T_0$, $T_0 + 1$, and $T_0 + 2$). SHAP results revealed that the model focuses most on the energy characteristics at the three time points: $T_0$, $T_0 + 1$, and $T_0 + 2$. When $T_0$ and subsequent sampling points exhibit high energy values while earlier points (e.g., $T_0 - 2$) display relatively low energy, the model tends to classify the event as gas-bearing. This observed rearward shift in the energy centroid aligns with the instantaneous phase delay and energy envelope tailing characteristics induced by gas-bearing reservoirs, reflecting the physical process of kinetic energy dissipation as seismic waves propagate through fluid-saturated media.

The comprehensive SHAP analysis revealed that the MR-SSGAN establishes a logically rigorous decision criterion on the synthetic data. In the frequency domain, the model relies critically on energy resonance within specific sensitive frequency bands, while in the time domain, it identifies evolving patterns of energy distribution asymmetry. This process not only statistically validates the model's high accuracy but also geophysically confirms its decision logic aligns closely with tuning effects and wavefield attenuation patterns.

### 3.2. Results of real seismic data

To validate the practical applicability of the proposed framework, this study employed actual data from deep marine carbonate reservoirs of the Leikoupu Formation in the Triassic strata of the western Sichuan Basin, China. The target interval (T1–T2), buried at depths exceeding 5,000 m, exhibits strong heterogeneity, complex pore structures, and extremely thin interbeds due to multiple

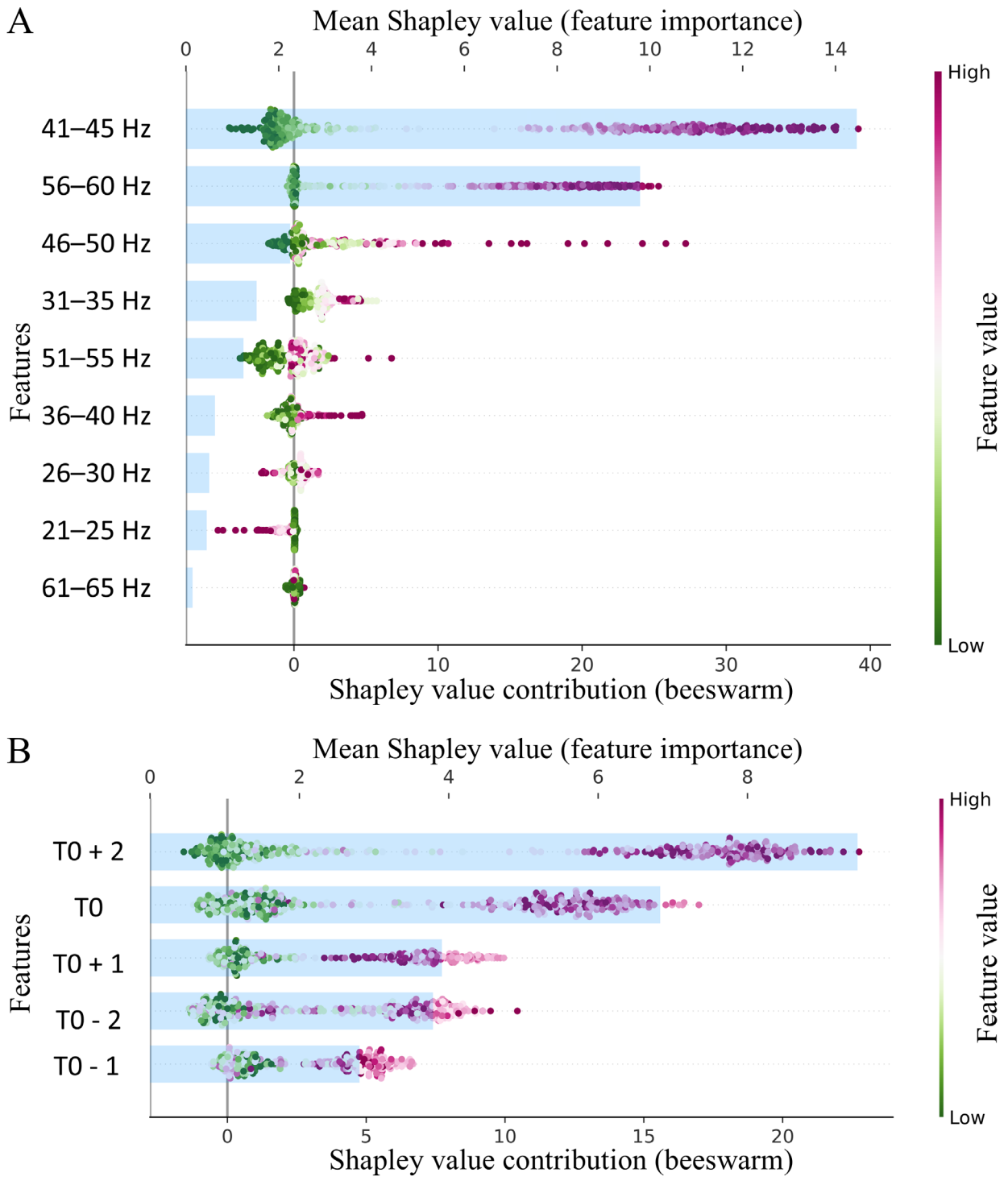



**Figure 8.** SHapley Additive exPlanations (SHAP) analysis of the manifold-regularized generative adversarial network on synthetic data. Each point represents a single feature from one sample, with purple and green corresponding to high and low feature values, respectively. The *x*-axis position denotes the SHAP value for that feature (positive values indicate promotion of gas-bearing predictions, whereas negative values indicate promotion of gas-free predictions). The bar lengths on the left intuitively reflect the average absolute SHAP value for each feature, indicating its importance. On the frequency dimension, the frequency axis of the TF plot is divided into bands at 5 Hz intervals.

episodes of tectonic deformation. This has resulted in severe absorption attenuation of seismic signals and an extremely low SNR environment. Furthermore, complex lateral facies transitions and uneven gas–water distribution pose significant challenges to gas content prediction due to non-uniqueness.

A representative 2D seismic profile (Figure 9) was selected, traversing three key calibration wells: Well A (gas-poor), Well B (gas-rich), and Well C (gas-bearing). From top to bottom, four main reflection horizons were identified within the profile: T1 (weathered crust at the top of the Leikoupo Formation), T2 (top of the fourth member of the Leikoupo Formation), T3 (top of the third member of the Leikoupo Formation), and T4 (bottom of the third member of the Leikoupo Formation). Intelligent detection focused specifically on gas content within the T1–T2 interval, aiming to validate the model's capability for detailed characterization of actual gas presence within complex geological settings.

#### 3.2.1. Feature response in real data

This study conducted TF characterization analysis on seismic records from the bypass Well B. The results (Figure 10)demonstrated high consistency with the synthetic data experiments: STFT remained constrained by the Heisenberg uncertainty principle, leading to TF energy dispersion, whereas FSST and WSST significantly enhanced feature concentration and resolution through their phase rearrangement mechanisms. Further comparison of the 15 Hz frequency slices (Figure 11) revealed that the energy clusters identified by traditional STFT and CWT in the T1–T2 interval were primarily distributed between Wells A and B, failing to effectively characterize the known gas-bearing zone near Well C.

Following application of SST, the WSST frequency slice (Figure 11C) revealed scarcely any pronounced zones of high energy concentration within the target stratum. While the FSST frequency slice (Figure 11D) demonstrates improved local energy contrast, it still lacked physical responses highly consistent with drilling conditions at the critical Well B and Well C intervals. This phenomenon clearly demonstrates that, for deep, strongly heterogeneous thin reservoirs in western Sichuan, conventional spectral decomposition properties remain constrained by seismic multi-solutions. Consequently, even with advanced synchrosqueezing operators, precise localization of gas-bearing sweet spots remains challenging.

#### 3.2.2. Accuracy and section evaluation

Given the complexity of seismic data from the field area, this study employed a 2D sliding TF window technique consistent with the synthetic experiments to construct samples, systematically evaluating the predictive performance of MR-SSGAN under sparse labeling conditions.

During feature extraction, FSST was applied to perform TF transformations on each raw seismic trace. Given the low dominant frequency (approximately 20 Hz) characteristic of deep reflection signals, the frequency truncation window was set to 1–45 Hz to ensure coverage of the core energy range spanning 5th–95th percentiles of most samples. The time window comprised the target sampling point flanked by two preceding and two following points. Given the 500 Hz sampling rate of the field data, the resulting feature slice dimensions remained 45 × 5, corresponding to a time-domain length of 10 ms.

To scientifically evaluate the model's generalization capability in strongly heterogeneous environments, a neighborhood $N_{W_i}(p)$ of $p$ units around well $W_i$ was defined in the coordinate system composed of xline and

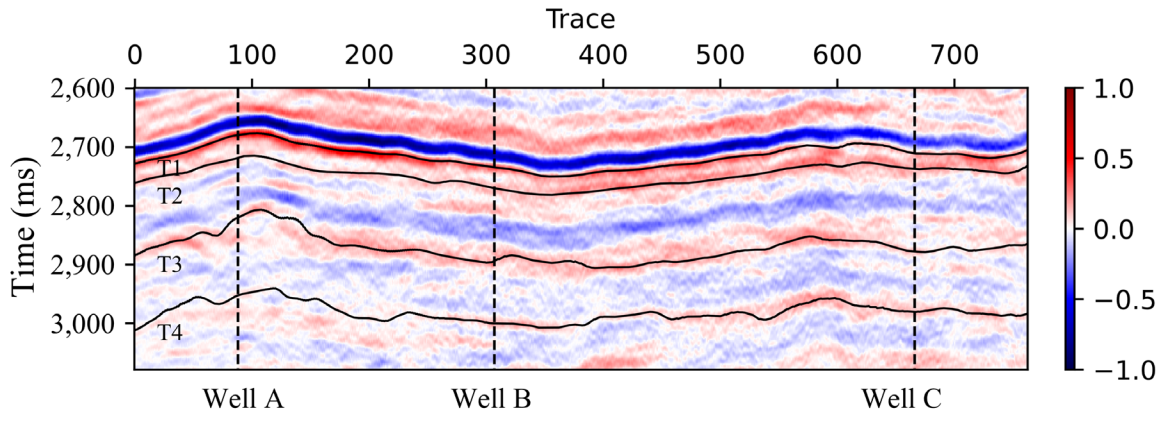



**Figure 9.** A post-stack seismic profile from the field area

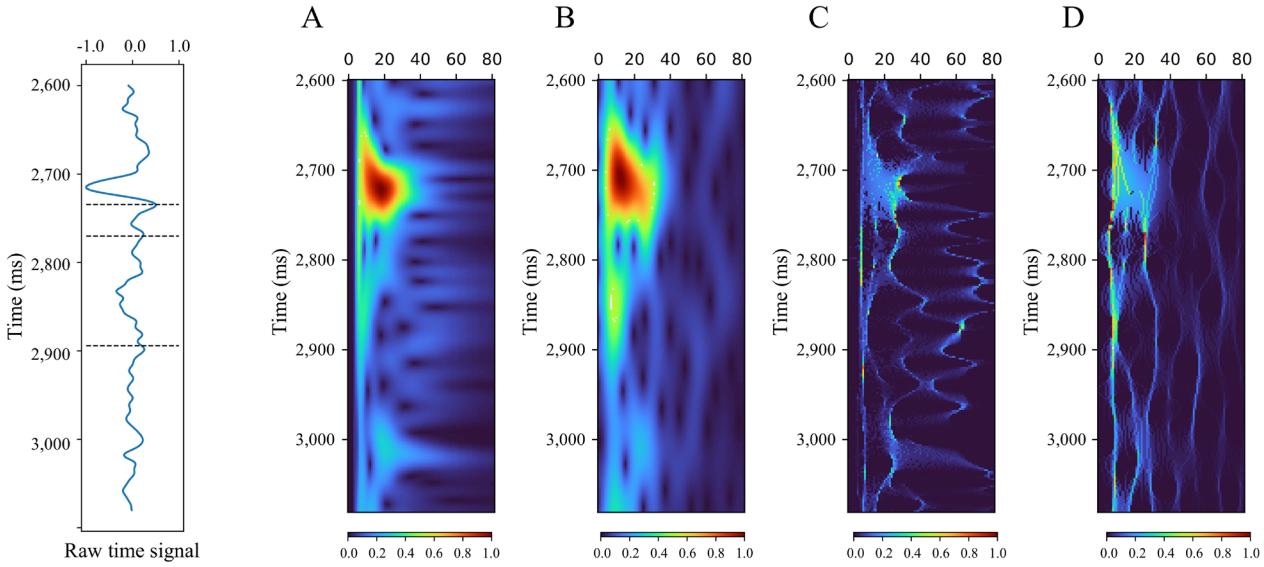



**Figure 10.** The original time-domain signal of the well-side trace at Well B and its corresponding time–frequency analysis results: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

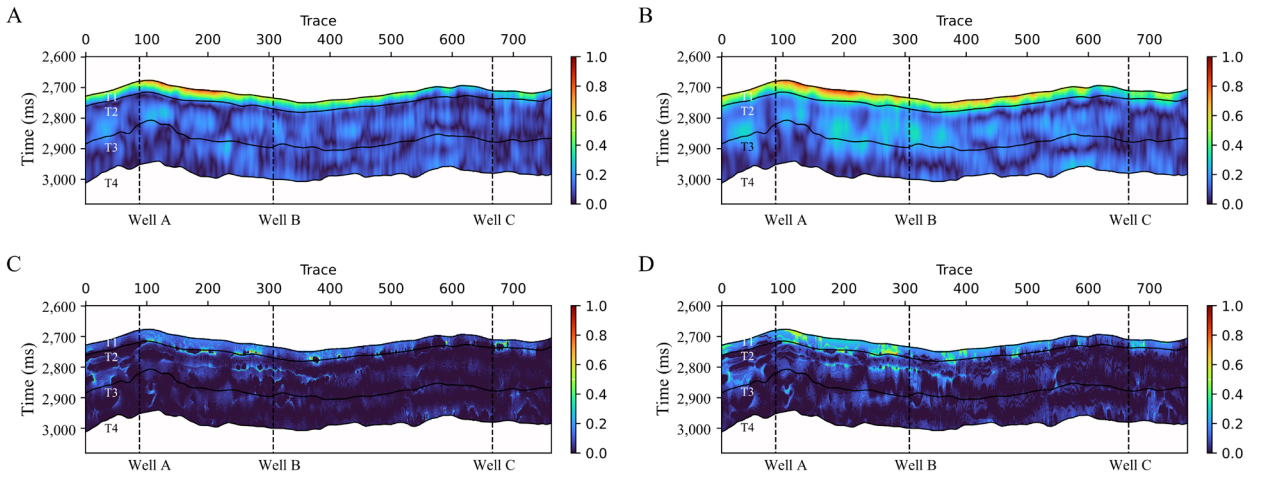



**Figure 11.** 15 Hz frequency slices of real data: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

inline:

$$\left\{(xl,il)\in R^2\,\middle|\,p\le\sqrt{(xl-xl_w)^2-(il-il_w)^2}<p+1\right\} \tag{13}$$

where the well location $(xl_{Wi},il_{Wi})$ is represented by $N_{Wi}(0)$. Besides the well location, the neighborhood one unit away from the well is represented by $N_{Wi}(1)$, which includes eight seismic traces, namely $(xline\pm1,inline)$, $(xline\pm1,inline-1)$, $(xline\pm1,inline+1)$, and $(xline,inline\pm1)$. Wells B and C were used as training wells, with labeled samples selected from the well-side traces. Unlabeled training samples, validation samples, and test samples were derived from seismic traces within a neighborhood of three units outside the training well locations. The specific sample quantity allocation was as follows: labeled training samples $L_{real}=40$, unlabeled training samples $U_{real}=800$, test samples $T_{real}=40$, and the number of validation set samples was approximately 10% of the total training data. The learning rate, batch size, and early stopping strategy (patience = 20) for model training were generally consistent with the synthetic data experiments.

Table 3 presents a comparative analysis of different

model architectures on real seismic datasets (averaged across three randomized experiments). Under conditions of extremely sparse labeling, the semi-supervised strategy demonstrates significant advantages: SSGAN achieves an 18.8% improvement in accuracy (with an F1 score increase of 0.200) over the purely supervised GAN, robustly demonstrating the auxiliary role of unlabeled data. Building upon this, the introduction of manifold regularization in MR-SSGAN yields an additional 5.0% gain in accuracy. This phenomenon aligns closely with the trend observed in the synthetic experiments, further validating that manifold constraints effectively suppress prediction fluctuations caused by noise in the field area by enhancing model smoothness on the data manifold.

To investigate the coupled effects of feature quality and semi-supervised architecture, the classification performance was compared using four TF representations as input (see Table 4). Quantitative results confirm that the MR-SSGAN model trained with FSST significantly outperforms other approaches in both accuracy and F1 score. Prediction accuracy based on STFT, CWT, and WSST exhibits a substantial decline, primarily due to spectral leakage or spatiotemporal blurring during the processing of deep, strongly inhomogeneous seismic signals.

To systematically evaluate the impact of different TF inputs on model predictive performance, Figure 12 presents the predicted probability profiles of MR-SSGAN trained using four TF representations.

First, predictions based on CWT (Figure 12A) and STFT (Figure 12B) exhibit significant deviations from known geological interpretations. Although CWT possesses multiscale analysis capabilities, its energy distribution becomes excessively dispersed when applied to real data with low SNR. This resulted in missed detections within the critical gas-bearing zones of Wells B and C, while simultaneously generating false continuous high-probability bands in the gas-depleted zone of Well A. The STFT, constrained by the inherent resolution limitation due to fixed window lengths, produces blurred boundaries for predicted gas-bearing zones and generates large-scale artefactual anomalies within non-target intervals (T2–T3 sections). These observations indicate that traditional TF transforms were unable to effectively mitigate TF blurring caused by the Heisenberg uncertainty principle, rendering them incapable of providing robust discriminative features for models under complex background noise.

Second, although synchrosqueezing technology aims to enhance energy concentration, the WSST-based model (Figure 12C) exhibits suboptimal performance in practical applications. The predicted profile exhibits severe fragmentation, poor lateral continuity, and significant misclassification in the middle–lower section. This is because, when processing complex wavefields, the energy reassignment operator of WSST is highly sensitive to strong noise interference, thereby compromising the physical consistency of the input features. This further indicates that, within the strongly heterogeneous conditions of the field area, an algorithm's robustness to noise is equally critical to feature aggregation.

**Table 3. Performance comparison of different models based on Fourier-based synchrosqueezing transform data (real data)**

| Network architecture | Number of labeled samples | Performance | | Gain vs. SGAN | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| SGAN | $1 \times L_{real}$ | 61.3 | 0.613 | - | - |
| SSGAN | $1 \times L_{real}$ | 80.0 | 0.813 | 18.8 | 0.200 |
| MR-SSGAN | $1 \times L_{real}$ | 85.0 | 0.885 | 23.8 | 0.272 |

Abbreviations: MR: Manifold-regularized; SGAN: Supervised generative adversarial network; SSGAN: Semi-supervised generative adversarial network.

**Table 4. Performance comparison of different time–frequency data based on MR-SSGAN (real data)**

| Input | Number of labeled samples | Performance | | Loss vs. FSST | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| CWT | $1 \times L_{real}$ | 75.0 | 0.767 | 10.0 | 0.118 |
| STFT | $1 \times L_{real}$ | 77.9 | 0.784 | 7.1 | 0.100 |
| WSST | $1 \times L_{real}$ | 70.6 | 0.707 | 14.4 | 0.177 |
| FSST | $1 \times L_{real}$ | 85.0 | 0.885 | - | - |

Abbreviations: CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; SGAN: Supervised generative adversarial network; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

Third, the FSST-based MR-SSGAN architecture (Figure 12D), in contrast, exhibits superior predictive performance, with results highly consistent with drilling calibration. This model not only accurately captured the gas anomalies in the target intervals of Wells B and C, presenting clear and laterally coherent high-probability responses, but also precisely depicted the gas-depleted conditions near Well A. The predicted reservoir morphology is well-defined, with thickness characterization better aligned with the geological development features of thin interbedded formations, demonstrating exceptionally high vertical and spatial resolution.

In summary, the experimental results further validate the synergistic effect between FSST and MR-SSGAN. FSST provides a highly focused and more noise-resistant feature benchmark through energy reassignment, while MR-SSGAN establishes a stable smoothing constraint in the feature space via manifold regularization. The deep integration of this high-quality input and robust learning paradigm enables the architecture to effectively overcome seismic ambiguity. Even under the constraint of extremely sparse labels, it achieves high-reliability detection of gas-bearing sweet spots in deep thin reservoirs in western Sichuan, providing effective technical support for the intelligent evaluation of complex hydrocarbon reservoirs.

### *3.2.3. SHapley Additive exPlanations analysis on real data*

SHapley Additive exPlanations analysis applied to field seismic data (Figure 13) further elucidates the decision logic of MR-SSGAN within complex geological environments, revealing attribution patterns markedly distinct from those observed in synthetic data.

Along the frequency dimension (Figure 13A), SHAP attribution results maintain high consistency with classical gas-bearing response theory. The feature importance bar chart indicates that the model primarily relies on information within the 6–30 Hz frequency band for decision-making, with the 6–10 Hz low-frequency range contributing most significantly. The scatter plot further reveals that energy enhancement (high feature values) in this low-frequency band is positively correlated with gas-bearing discrimination, accurately capturing the typical low-frequency shadow feature. In contrast, within the 11–30 Hz mid-to-high-frequency range, energy attenuation (low feature values) contributes positive Shapley values to gas-bearing prediction. This combined signature of low-frequency resonance and high-frequency absorption physically corroborates the expected dynamic evolution of seismic waves as they propagate through gas-bearing porous media. Compared to the mid-to-high-frequency

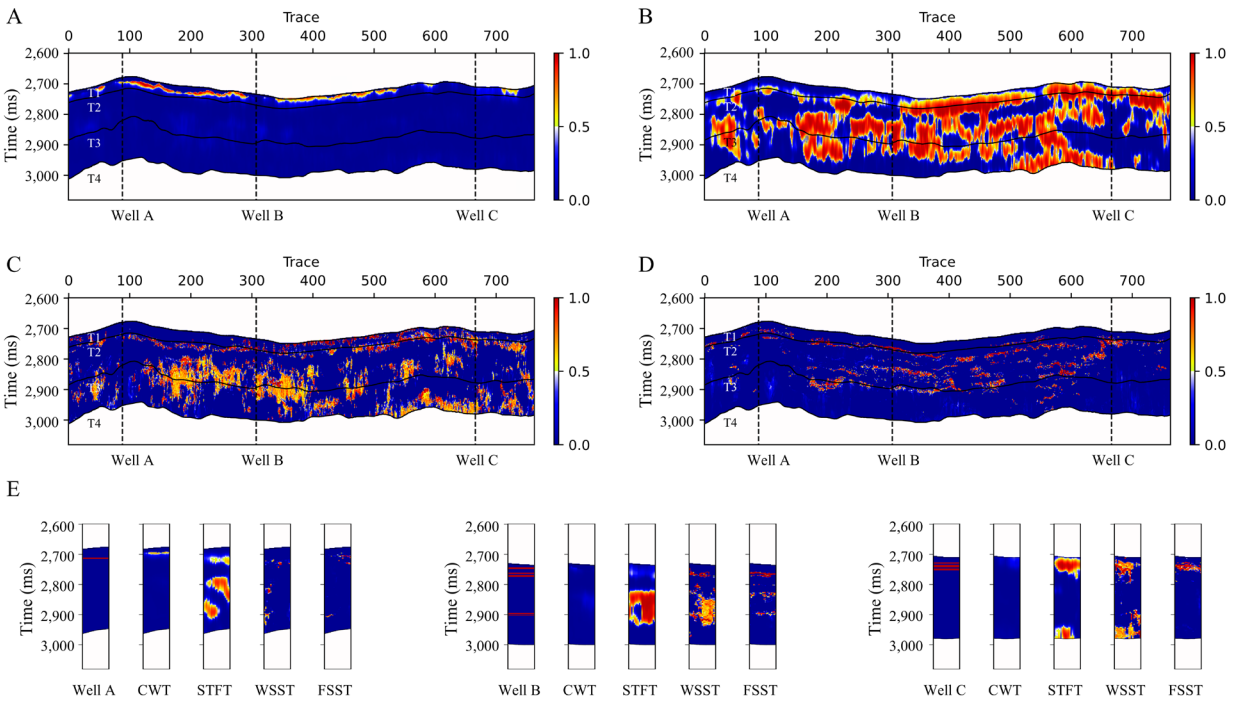



**Figure 12.** Gas-bearing reservoir prediction probability of the manifold-regularized generative adversarial network based on different time–frequency data (input is real seismic data): (A) CWT, (B) STFT, (Cc) WSST, and (D) FSST. (E) Comparison of labels versus predictions across time–frequency inputs. Abbreviations: CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; SGAN: Supervised generative adversarial network; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

enhancement pattern dominated by thin-layer tuning effects in synthetic data, the analysis results from real data align more closely with classical wavefield attenuation theory, reflecting the intelligent discriminative adaptation of the deep learning framework when addressing idealized models versus complex real-world environments.

In the temporal dimension (Figure 13B), the model exhibits sampling point dependency patterns distinct from synthetic data. Analysis indicates that the present time $T_0$ carries the highest importance weight, followed by $T_0 - 2$ and $T_0 + 2$. The model identifies a specific temporal evolution pattern: samples exhibiting localized energy enhancement at $T_0 - 2$ followed by energy attenuation at $T_0$ and $T_0 + 2$ are more readily identified as gas-bearing. This complex response pattern likely originates from the intricate wavefield interference and phase delay phenomena within deep, thin interbedded reservoirs. Compared to the simple energy center shift observed in synthetic data, the SHAP attribution patterns in real data reveal how actual stratigraphic absorption characteristics modulate the waveform envelope morphology.

The comprehensive SHAP attribution analysis demonstrates that the MR-SSGAN model not only achieves statistically high-precision predictions but also captures seismically plausible phenomena. The model adaptively adjusts its decision logic based on data context—transitioning from tuning-effect–driven behavior in synthetic data to absorption-attenuation–driven behavior in real data. This cross-dataset consistency in decision-making, coupled with transparent physical logic, effectively rules out the possibility of model classification arising from overfitting noise. Consequently, it significantly enhances the reliability and guidance value of artificial intelligence technology in high-risk decision-making within oil and gas exploration.

## 4. Discussion

### 4.1 Selection and parameter impact of synchrosqueezing transform

During the feature characterization phase, experimental results (Tables 2 and 4) show that FSST produces downstream models with more discriminative input features than WSST. Although the wavelet-based WSST has advantages in multiscale analysis, its energy rearrangement operator is sensitive to random noise when processing seismic signals with a low SNR. This results in fractured TF ridges and energy artifacts. In contrast, FSST integrates the DPSS window function, which has greater sidelobe suppression capabilities, resulting in a relatively robust analysis framework. This design successfully mitigates random disturbances during energy compression, guaranteeing that extracted gas-bearing anomaly responses are statistically stable. While this study provides initial validation of the chosen window parameters, future work should involve comprehensive sensitivity analyses to develop tailored optimization strategies for SST, targeting specific applications such as ultra-deep or tight gas reservoirs. This will further advance the capabilities of high-resolution TF characterization.

### 4.2. Critical role of time–frequency coupled features

Seismic signals are inherently non-stationary, complex wave processes. Traditional 1D attributes (e.g., instantaneous frequency or amplitude) represent time and frequency information separately, resulting in the loss of rich context-coupled characteristics. The comparative experiment (Figure 14) clearly illustrates this limitation: model predictions based solely on 1D instantaneous frequency exhibit pronounced fragmentation, lack geological lateral continuity, and struggle to characterize the fine boundaries of thin reservoirs. In contrast, 2D TF slices simultaneously capture complete spectral evolution patterns at specific moments and within their adjacent temporal windows. SHAP analysis (Figure 13B) further confirms, from an attribution perspective, that the model's gas-bearing classification heavily depends on the energy distribution ratio within the narrow time window from $T_0 - 2$ to $T_0 + 2$. This cross-dimensional spatiotemporal coupling characteristic cannot be expressed by 1D attributes and constitutes the key physical prerequisite for achieving high-precision identification of complex thin reservoirs.

### 4.3. Comparison of learning paradigms

When addressing the challenge of sparse labeling, the proposed MR-SSGAN demonstrates greater endogeneity and robustness compared to data augmentation, which depends on manually generated samples[31,32] or transfer learning[11,33,34], which relies on external source domain knowledge. Although the data augmentation strategy (Figure 15A) enhances coverage, its predicted profiles are often accompanied by false positive indications of questionable geological plausibility and are highly dependent on the quality of generated pseudo-samples. Transfer learning (Figure 15B) focuses on sweet-spot regions, yet suffers from weak continuity representation in inter-well areas due to domain differences between source and target domains.

In contrast, MR-SSGAN achieves end-to-end knowledge extraction by directly mining the distributional manifold of vast unlabeled data within the target field area. Its predictions strike a favorable balance between smoothness in macro-trends and focus on local details (Figure 12D).

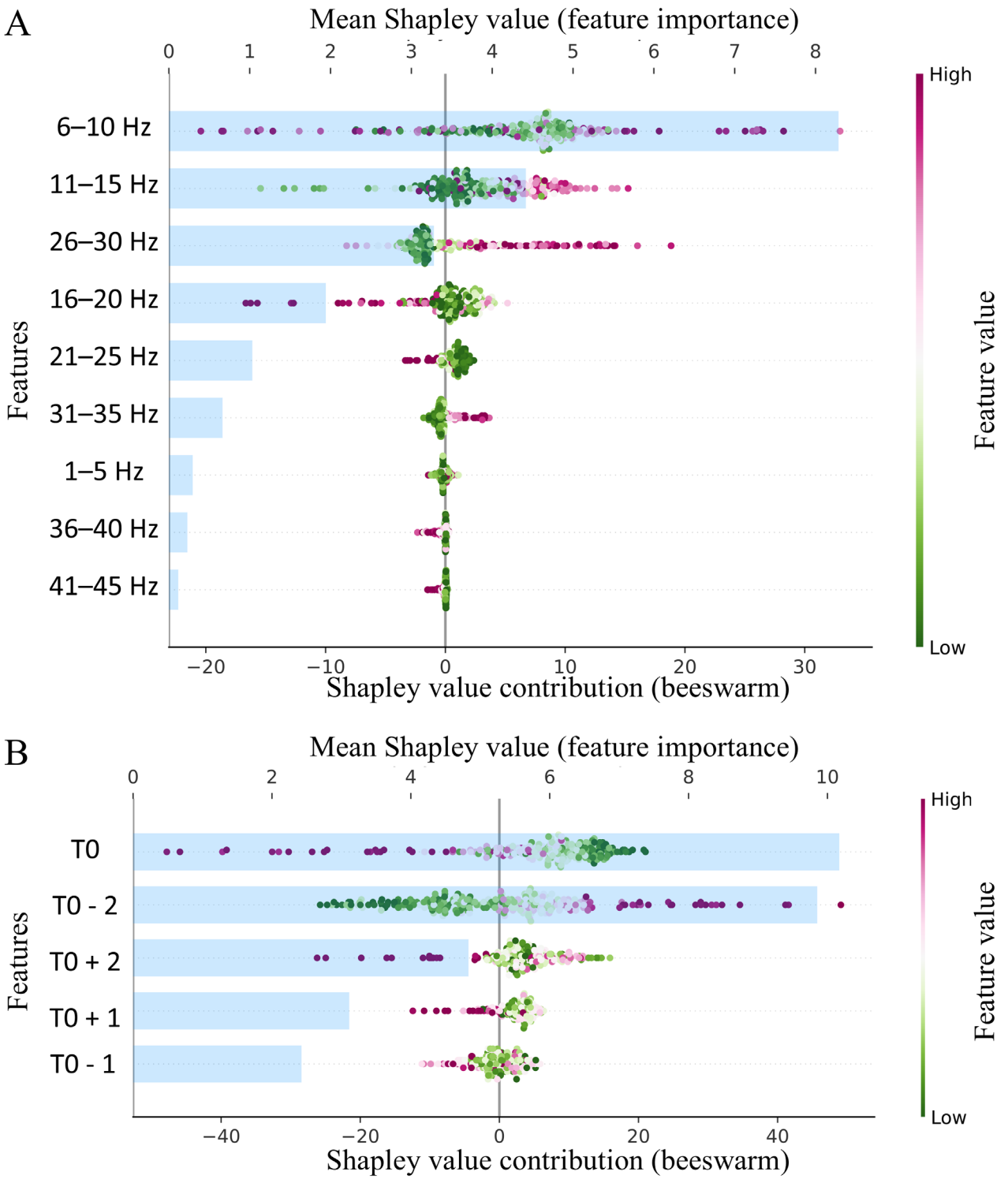



**Figure 13.** SHapley Additive exPlanations analysis of the manifold-regularized generative adversarial network on real seismic data

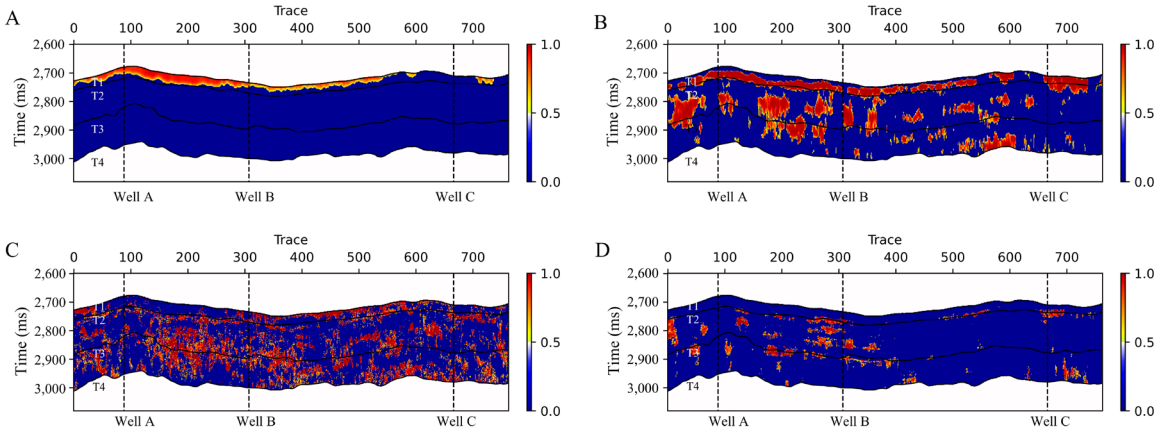



**Figure 14.** Gas-bearing reservoir prediction probability of the manifold-regularized generative adversarial network based on different one-dimensional instantaneous frequency inputs from real seismic data: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

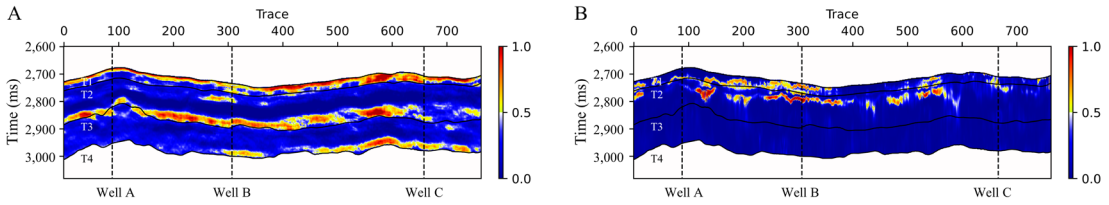



**Figure 15.** Gas-bearing reservoir prediction probability using (A) a dual-branch spatio-temporal fusion model with data augmentation and (B) an efficient transformer model with transfer learning

This semi-supervised paradigm, which directly acquires prior information from the target domain's topological geometric features, not only reduces reliance on external training datasets but also enhances the algorithm's generalizability in highly heterogeneous environments. Future exploration of embedding small quantities of high-quality augmented data into the semi-supervised training process holds promise for further improving the intelligent evaluation efficiency of complex reservoirs.

## 5. Conclusion

This study proposes an intelligent detection framework that incorporates the STFT-based FSST, MR-SSGAN, and SHAP analysis to address core challenges in gas content identification within complex reservoirs, namely scarce labeled samples, significant environmental noise interference, and opaque model decision-making. This paradigm shows strong prediction advantages in both synthetic models and real seismic data. Under extreme label scarcity (5% labeled samples), the SSL technique and manifold-regularization constraint improved accuracy by 18.8% and 5.0%, respectively (with simultaneous significant gains in F1 scores). Ablation experiments show that 2D TF coupled information better describes transient dynamic anomaly responses in thin reservoirs than 1D transient features. Furthermore, unlike transfer learning, which relies on external previous information or data augmentation using artificial rules, the proposed semi-supervised architecture recovers latent distributional patterns directly from real seismic data. This demonstrates higher generalization capacity when dealing with low SNR settings, such as deep marine reservoirs in the Sichuan Basin. Further SHAP analysis demonstrates that the MR-SSGAN model adapts its decision logic based on data context, shifting from tuning-effect–driven in synthetic data to absorption-attenuation–driven in real data. This capability not only verifies the model's ability to capture true physical principles but also aligns its decision-making process with conventional rock physics theories. This considerably improves the dependability of data-driven strategies in high-risk decision-making for oil and gas exploration.

Looking ahead, research may deepen in two directions: first, introducing automatic optimization algorithms—

such as particle swarm optimization or Bayesian optimization—for adaptive tuning of key parameters to further enhance the framework's generalizability; second, integrating advanced architectures like Transformers with multi-source physical information to uncover deeper sensitivity patterns, ultimately constructing a more precise and reliable intelligent evaluation system for hydrocarbon exploration.

## Acknowledgments

None.

## Funding

This work was supported by the National Natural Science Foundation of China (Grant 42330813 and Grant 42030812), and the Natural Science Foundation of Sichuan Province (Grant 2026NSFSC1138).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Shuying Ma, Junxing Cao, Rong Wang
*Formal analysis:* Jun Wang, Lingsen Zhao, Hong Li, Xin Tang
*Investigation:* Junxing Cao, Xudong Jiang
*Methodology:* Shuying Ma
*Writing–original draft:* Shuying Ma
*Writing–review & editing:* Shuying Ma, Junxing Cao, Rong Wang

## Availability of data

Data supporting the findings of this study are available from the corresponding authors upon reasonable request.

## References

1. Hammond AL. Bright Spot: Better Seismological Indicators of Gas and Oil. *Science*. 1974;185(4150):515-517.

   doi: 10.1126/science.185.4150.515

2. Brown AR. Dim spots: Opportunity for future hydrocarbon discoveries? *Lead Edge*. 2012;31(6):682-683.

   doi: 10.1190/tle31060682.1

3. Backus MM, Chen RL. Flat spot exploration. *Geophys Prospect*. 1975;23(3):533-577.

   doi: 10.1111/j.1365-2478.1975.tb01547.x

4. Korneev VA, Goloshubin GM, Daley TM, Silin DB. Seismic low-frequency effects in monitoring fluid-saturated reservoirs. *Geophysics*. 2004;69(2):522-532.

   doi: 10.1190/1.1707072

5. Silin DB, Korneev VA, Goloshubin GM, Patzek TW. *A Hydrologic View on Biot's Theory of Poroelasticity*. Office of Scientific and Technical Information (OSTI); 2004.

   doi: 10.2172/822181

6. Mitchell JT, Derzhi N, Lichman E, Lanning EN. Energy absorption analysis: A case study. In: *SEG Technical Program Expanded Abstracts 1996*. 66th SEG Annual International Meeting; January 1, 1996; Denver, Colorado, USA. ASME; 1996:1785-1788.

   doi: 10.1190/1.1826480

7. Santos LF, Silva RMGE, Gattass M, Silva AC. Direct hydrocarbon indicators based on long short-term memory neural network. In: *SEG Technical Program Expanded Abstracts 2019*. 89th SEG Annual International Meeting; August 1, 2019; San Antonio, Texas, USA. ASME; 2019:2373-2377.

   doi: 10.1190/segam2019-3215628.1

8. Júnior DAD, Batista da Cruz L, Otávio Bandeira Diniz J, *et al*. Detection of potential gas accumulations in 2D seismic images using spatio-temporal, PSO, and convolutional LSTM approaches. *Expert Syst Appl*. 2023;215:119337.

   doi: 10.1016/j.eswa.2022.119337

9. Ma S, Cao J, Liu Z, Jiang X, Su Z, Xue Y. Gas-bearing prediction of deep reservoir based on DNN embeddings. *Front Earth Sci*. 2023;11.

   doi: 10.3389/feart.2023.1117797

10. Jiang X, Cao J, Zu S, Xu H, Wang J. Detection of hidden reservoirs under strong shielding based on bi-dimensional empirical mode decomposition and the Teager–Kaiser operator. *Geophys Prospect*. 2021;69(5):1086-1101.

    doi: 10.1111/1365-2478.13073

11. Gao J, Song Z, Gui J, Yuan S. Gas-Bearing Prediction Using Transfer Learning and CNNs: An Application to a Deep Tight Dolomite Reservoir. *IEEE Geosci Remote Sensing Lett*. 2022;19:1-5.

    doi: 10.1109/lgrs.2020.3035568

12. Yang J, Lin N, Zhang K, Zhang D, Wang D, Zhang J. Variable Selection and Oil-Gas-Bearing Evaluation Using Multicomponent Seismic Data and AMPSO-ML Approaches. *SPE J*. 2023;28(03):1183-1201.

    doi: 10.2118/214298-pa

13. Yang JQ, Lin NT, Zhang K, Cui Y, Fu C, Zhang D. Deep learning CNN-APSO-LSSVM hybrid fusion model for feature optimization and gas-bearing prediction. *Pet Sci*. 2024;21(4):2329-2344.

    doi: 10.1016/j.petsci.2024.02.012

14. Yang X, Song Z, King I, Xu Z. A Survey on Deep Semi-Supervised Learning. *IEEE Trans Knowl Data Eng*. 2023;35(9):8934-8954.

doi: 10.1109/tkde.2022.3220219

15. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, *et al*. Generative adversarial networks. *Commun ACM*. 2020;63(11):139-144.

doi: 10.1145/3422622

16. Odena A. Semi-supervised learning with generative adversarial networks. *arXiv*. Preprint online 2016.

doi: 10.48550/arXiv.1606.01583

17. Liu M, Jervis M, Li W, Nivlet P. Seismic facies classification using supervised convolutional neural networks and semisupervised generative adversarial networks. *Geophysics*. 2020;85(4):O47-O58.

doi: 10.1190/geo2019-0627.1

18. Song L, Yin X, Zhang R, Li J, Zhang J, Li J. Reservoir Discrimination Based on Physic-Informed Semi-Supervised Learning. *IEEE Trans Geosci Remote Sens*. 2024;62:1-14.

doi: 10.1109/tgrs.2024.3409578

19. Song L, Yin X, Yin L. Reservoir Lithology Identification Based on Improved Adversarial Learning. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

doi: 10.1109/lgrs.2023.3281545

20. Daubechies I, Lu J, Wu HT. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl Comput Harmon Anal.* 2011;30(2):243-261.

doi: 10.1016/j.acha.2010.08.002

21. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Guyon I, Von Luxburg U, *et al*., editors. *Advances in Neural Information Processing Systems 30*. 31st International Conference on Neural Information Processing Systems (NIPS 2017); December 4–9, 2017; Long Beach, CA, USA. Curran Associates Inc; 2017:4766-4777.

22. Huang Z lai, Zhang J, Zhao T hu, Sun Y. Synchrosqueezing S-Transform and Its Application in Seismic Spectral Decomposition. *IEEE Trans Geosci Remote Sens*. 2016;54(2):817-825.

doi: 10.1109/tgrs.2015.2466660

23. Mao X. A Concentrated Time-Frequency Method for Reservoir Detection Using Adaptive Synchrosqueezing Transform. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/lgrs.2022.3160930

24. Belkin M, Niyogi P, Sindhwani V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. J Mach Learn Res. 2006;7(85):2399–2434.

25. Ni Y, Koniusz P, Hartley R, Nock R. Manifold Learning Benefits GANs. In: *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). June 19–24, 2022; New Orleans, Louisiana. IEEE; 2022:11255-11264.

doi: 10.1109/cvpr52688.2022.01098

26. Lecouat B, Foo CS, Zenati H, Chandrasekhar VR. Semi-Supervised Learning with GANs: Revisiting Manifold Regularization. *arXiv*. Preprint posted online May 23, 2018. Last accessed November 4, 2024.

doi: 10.48550/arXiv.1805.08957

27. Oberlin T, Meignen S, Perrier V. The fourier-based synchrosqueezing transform. In: *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*). May 4–9, 2014; Florence, Italy. IEEE; 2014:315-319.

doi: 10.1109/icassp.2014.6853609

28. Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X. Improved Techniques for Training GANs. *arXiv*. Preprint posted online June 10, 2016. Last accessed November 1, 2024.

doi: 10.48550/arXiv.1606.03498

29. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: *Proceedings of the 32nd International Conference on Machine Learning*. International Conference on Machine Learning; July 7-9 2015; Lille, France. PMLR; 2015;37:448-456.

30. Martin GS, Marfurt KJ, Larsen S. Marmousi-2: An updated model for the investigation of AVO in structurally complex areas. In: *SEG Technical Program Expanded Abstracts 2002*. 72th SEG Annual International Meeting; January 1, 2002; Salt Lake City, Utah, USA. ASME; 2002:1979-1982.

doi: 10.1190/1.1817083

31. Ma S, Cao J. A Two-Branch Neural Network for Gas-Bearing Prediction Using Latent Space Adaptation for Data Augmentation—An Application for Deep Carbonate Reservoirs. *IEEE Geosci Remote Sens Lett*. 2024;21:1-5.

doi: 10.1109/lgrs.2024.3436836

32. Zhang K, Lin N, Yang J, Zhang D, Cui Y, Jin Z. An Intelligent Approach for Gas Reservoir Identification and Structural Evaluation by ANN and Viterbi Algorithm—A Case Study From the Xujiahe Formation, Western Sichuan Depression, China. *IEEE Trans Geosci Remote Sens*. 2023;61:1-12.

doi: 10.1109/tgrs.2023.3247183

33. Song Z, Li S, He S, Yuan S, Wang S. Gas-Bearing Prediction of Tight Sandstone Reservoir Using Semi-Supervised Learning and Transfer Learning. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/lgrs.2022.3177314

34. Ma S, Cao J. An Efficient Transformer Model Enhanced by S-Transform and Transfer Learning for Predicting Gas Distribution in Deeply Buried Reservoirs. *IEEE Trans Geosci Remote Sens*. 2025;63:1-14.

doi: 10.1109/tgrs.2025.3558731

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Two-way acoustic wave-equation-based wavefield depth extrapolation for imaging in vertically transversely isotropic media

**Shanzheng Hu[1,2], Pengyuan Sun[1,2], Shuqin Li[1,2], Feng Hu[1,2], Min Guan[1,2], and Jiachun You[3]***

[1]BGP Inc., China National Petroleum Corporation, Zhuozhou, Hebei, China

[2]National Engineering Research Center of Oil and Gas Exploration Computer Software, China National Petroleum Corporation, Zhuozhou, Hebei, China

[3]Department of Geophysics, Chengdu University of Technology, Chengdu, Sichuan, China

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

## Abstract

Anisotropy is a prevalent characteristic of the Earth's subsurface, giving rise to more complex wavefield behavior than in purely acoustic media. Seismic data acquired in such anisotropic environments pose significant challenges for conventional acoustic imaging methods that rely on isotropic assumptions and acoustic approximations, often leading to inaccurate imaging of complex geologic structures. To address these challenges, we propose a wavefield depth extrapolation method grounded in the two-way acoustic anisotropic wave equation, specifically tailored for imaging in vertically transversely isotropic (VTI) media. Theoretically, the proposed method extends the two-way wave-equation-based wavefield extrapolation framework from acoustic media to anisotropic VTI media. To further improve image quality, we used a frequency–wavenumber-domain approach to remove pseudo-S-wave energy, enhancing the clarity of P-wave descriptions and reducing imaging artifacts. We evaluated the effectiveness of the proposed approach through several numerical tests, including experiments on a VTI three-layer model, the Hess VTI model, and the Marmousi VTI model. These benchmarks demonstrated that our method consistently outperformed conventional acoustic schemes designed for isotropic media, yielding clearer, higher-resolution structural images. In addition, we applied both the acoustic and the VTI-based migration methods to a real seismic dataset. The VTI-based migration provided a clearer, more continuous image of the deep structures than the acoustic migration method. The numerical experiments and real data application demonstrated that the proposed VTI imaging strategy is both practical and effective. More importantly, the findings underscore the need to explicitly account for anisotropic parameters in seismic imaging workflows to achieve geologically consistent and reliable interpretations in VTI-dominated regions.

***Keywords:*** Vertically transversely isotropic media; Two-way acoustic wave equation; Wavefield depth extrapolation; Depth migration

***Corresponding author:**
Jiachun You
(youjiachun@cdut.edu.cn)

**Citation:** Hu S, Sun P, Li S, Hu F, Guan M, You J. Two-way acoustic wave-equation-based wavefield depth extrapolation for imaging in vertically transversely isotropic media. *J Seismic Explor.* 2026;35(2):025500124. doi: 10.36922/JSE025500124

**Received:** December 9, 2025

**Revised:** February 7, 2026

**Accepted:** February 7, 2026

**Published online:** April 27, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

In seismic exploration and geophysical imaging, the influence of medium anisotropy on wavefield propagation is non-negligible and complicated. Alkhalifah[1,2] demonstrated that velocity anisotropy is pervasive in subsurface formations, with vertically transversely isotropic (VTI) media being particularly dominant in practical exploration settings. Therefore, accurately accounting for VTI characteristics during seismic wavefield modeling, migration, and inversion is critical for improving imaging resolution and positioning accuracy. Over the past few decades, numerous numerical methods and imaging strategies developed for VTI media have been introduced, yielding significant progress. From a comprehensive perspective, Li and Qu[3] reviewed the development of seismic imaging, summarizing the transition from stacking- and migration-based methods to inversion-driven high-resolution approaches and highlighting the role of advanced techniques in complex media.

The accurate imaging of seismic data in VTI media remains a critical topic due to the prevalence of anisotropy in sedimentary basins. Duveneck *et al.*[4] derived a system of first-order acoustic wave equations specifically designed for VTI media and applied it to reverse-time migration (RTM), demonstrating reliable amplitude preservation and robustness against density variations. However, conventional acoustic approximations often suffer from instability, especially in the absence of shear-wave velocity. Bube *et al.*[5] highlighted that such approximations can lead to time-growing instabilities in second-order formulations, which necessitated the development of more stable wavefield propagation schemes. Zhou *et al.*[6] proposed a simplified anisotropic acoustic wave equation for VTI media by decomposing the fourth-order differential equation into a coupled lower-order system via an auxiliary function, facilitating implementation while retaining physical meaning.

To more effectively mitigate P/S wave crosstalk in VTI media, various wave-mode separation techniques have been developed. Yan and Sava[7] proposed pseudo-derivative operators constructed from polarization vectors obtained via the Christoffel equation, enabling effective elastic wavefield decomposition. Zhang and McMechan[8] extended this approach into the wavenumber domain, developing a robust vector decomposition technique for both 2D and 3D elastic wavefields. More recently, Yao *et al.*[9] proposed a pseudo-Helmholtz decomposition operator that leverages wavefront phase direction and eigensystem analysis, offering both computational efficiency and improved mode separation precision. Sun *et al.*[10] introduced a pure qP-wave RTM method using an elliptical decomposition vector equation, effectively addressing instability and crosstalk issues in VTI media. Shan *et al.*[11] introduced a least-squares Gaussian beam migration method for VTI media, addressing anisotropic effects and improving resolution and amplitude fidelity. Li *et al.*[12] proposed a generalized two-way phase shift operator for acoustic wave-equation depth migration, enhancing subsurface imaging in VTI media. Qin *et al.*[13] proposed a new pure quasi-P wave equation for VTI media, using optical flow to accurately calculate wave propagation directions and correct for anisotropic effects in RTM. Zhang *et al.*[14] extended RTM to VTI media with irregular surfaces, incorporating a boundary-fitted grid to improve imaging accuracy.

Regarding wavefield extrapolation and numerical modeling, low-rank symbol approximation methods have gained popularity. Fomel *et al.*[15] presented a decomposition method in the space-wavenumber domain, which greatly improved the computational efficiency of wavefield propagation in anisotropic media. Sun and Sun[16] further incorporated De Wolf approximation and VTI-generalized screen operators, mitigating aliasing artifacts and enhancing imaging accuracy for complex VTI structures. To account for attenuation effects, Mao *et al.*[17] formulated a viscoacoustic wave equation using complex dispersion relations, combined with the hybrid finite-difference and low-rank decomposition algorithm, accurately modeling qP-wave propagation and dissipation. Similarly, Liang *et al.*[18] developed a decoupled qP-wave equation specifically for VTI media, effectively suppressing artificial qSV-wave dispersion and thereby enhancing the numerical stability and accuracy of the wave-mode representation. Bai *et al.*[19] proposed a viscoelastic anisotropic wave equation for VTI media that decouples dissipation and dispersion, restoring time symmetry via adjusted dissipation terms to improve time-reversal imaging fidelity. Qu *et al.*[20] proposed a Q-compensated least-squares RTM method with velocity-anisotropy correction based on first-order velocity–pressure equations, demonstrating improved deep imaging resolution and amplitude preservation in VTI-attenuating media. Wang *et al.*[21] derived a fractional Laplacian viscoelastic wave equation for VTI-attenuating media, enabling accurate modeling of frequency-independent Q with high computational efficiency.

On the inversion front, advances in full-waveform and reflection-based techniques have enabled more reliable recovery of anisotropic parameters. Wu *et al.*[22] accelerated reflection waveform inversion in acoustic VTI media by integrating second-order adjoint-state methods with a matrix-free Gauss–Newton approach. Singh *et al.*[23] introduced geologically constrained elastic full-

waveform inversion that integrates rock-physics-informed regularization, significantly alleviating parameter coupling and underdetermination in VTI settings. In a probabilistic framework, Lang *et al.*[24] combined a five-parameter VTI reflection model with Bayesian Markov Chain Monte Carlo inversion, enhancing parameter resolution and uncertainty quantification. Moreover, Wang *et al.*[25] focused on multicomponent inversion using a modified SH–SH wave reflection model, yielding high-resolution estimates of anisotropic shear-wave properties. Kamath and Tsvankin[26] performed sensitivity analysis using Fréchet kernels, elucidating the directional radiation behavior of Thomsen parameters and revealing critical insights into wavefield parameter sensitivity and cross-coupling. Li *et al.*[27] analyzed RTM extended images for VTI media, showing η-related residual moveout is dip-dependent and δ errors cause defocusing only laterally, guiding anisotropic tomography. Luo *et al.*[28] proposed a hierarchical prestack inversion scheme for VTI media using exact reflection coefficients, thereby improving parameter estimation via a two-stage linear-nonlinear optimization. With the rapid advancement of deep learning, a new class of imaging approaches based on physics-informed neural networks (PINNs) has emerged in recent years. Song *et al.*[29] proposed employing PINNs to model the transmission of acoustic waves in VTI media.

To address the challenges of seismic imaging in anisotropic media, researchers have extensively worked on designing and improving one-way wave-equation depth-imaging techniques tailored to such complex media. Ristow[30] pioneered a finite-difference-based one-way depth-migration method tailored for transversely isotropic media. Building on this foundation, Han and Wu[31] introduced a dual-domain one-way propagator specifically designed for VTI media, enhancing wavefield extrapolation accuracy. Bale[32] proposed a migration technique based on the anisotropic acoustic wave equation, utilizing a phase-shift operator obtained from an adjusted dispersion relation. Fei and Liner[33] extended the hybrid Fourier–finite-difference (Fourier-FD) operator to a 3D depth migration framework, offering improved modeling capabilities for complex anisotropic structures. Bakker[34] developed a more stable one-way propagator for VTI media using a four-way splitting Fourier-FD method. Amazonas *et al.*[35] proposed a hybrid finite-difference operator based on complex Padé approximations, demonstrating enhanced stability for steeply dipping reflectors and effectively reducing migration artifacts. Salcedo *et al.*[36] optimized Padé coefficients to construct a wide-angle continuation operator with improved velocity dependence. Liu and Zhang[37] refined the operator coefficients in the wide-angle regime and incorporated a phase-shift mechanism to improve imaging precision in VTI media. Vyas *et al.*[38] explored high-resolution one-way wave propagation in tilted transversely isotropic media, thereby expanding the applicability of such methods. Alshuhail and Verschuur[39] proposed a robust migration strategy for VTI media that incorporates internal multiples through a generalized one-way phase-shift operator, thereby improving the reliability of anisotropic parameter estimation via joint migration inversion.

In the context of seismic imaging in anisotropic media, two-way wave-equation-based extrapolation has received increased attention due to its inherent ability to accurately simulate full wavefield behavior.[40–44] This advantage is particularly critical in VTI media, where strong anisotropy produces significant deviations in phase velocity, ray bending, and reflection moveout that standard acoustic or isotropic assumptions cannot adequately handle. Recent studies have shown that two-way propagators can improve the imaging of complex geological structures, particularly in settings with strong vertical velocity gradients or lateral heterogeneity.[45,46] These motivations collectively demonstrate that extending two-way extrapolation methods from isotropic to VTI media is not only theoretically desirable but also essential for achieving accurate and reliable imaging in anisotropic sedimentary basins.

Despite this progress, limited studies have further explored two-way acoustic wave-equation depth migration for VTI media. The contributions of this study can be summarized as follows. First, we extend the two-way acoustic wave-equation depth extrapolation framework to VTI media by deriving a first-order system that explicitly incorporates anisotropic coefficients, enabling accurate wavefield continuation in VTI environments. Second, we introduce a frequency–wavenumber (FK)-domain pseudo-S-wave suppression strategy that eliminates evanescent and spurious wave modes generated during two-way extrapolation, significantly improving the physical fidelity of the propagated P-wavefield. Third, inspired by recent advances in up–down wavefield decomposition, we formulate a separation operator tailored to the VTI two-way framework, enabling stable imaging conditions that reduce low-frequency artifacts. Finally, through comprehensive numerical experiments—including a three-layer model, the Hess model, the Marmousi model, and real land data—we demonstrate that the proposed method consistently outperforms conventional acoustic migration by producing clearer structures, enhanced reflector continuity, and reduced dispersion artifacts. These innovations collectively establish a robust and practical two-way VTI imaging methodology that advances the

current state of anisotropic seismic imaging.

## 2. Methodology

In heterogeneous media, numerical simulation of the elastic wave equation involves solving a coupled system (two components in 2D and three in 3D) relating to the vector elements of the wavefield. The computational cost of propagating each component is comparable to that in the acoustic case, but elastic modeling imposes additional demands on data handling, including the input, output, and storage of wavefield components alongside their corresponding elastic parameters. Unlike the acoustic formulation, which yields only compressional (P) waves, the elastic solution inherently includes both P- and shear (S) waves. This simultaneous propagation of multiple wave modes, while more complete physically, can be inefficient and less targeted for applications focused solely on P-wave modeling and imaging. Alkhalifah[2] introduced an acoustic wave equation specific to VTI media to address this problem, which can be expressed as follows:

$$\frac{\partial^2 p}{\partial t^2} = (1+2\eta)v_n^2\frac{\partial^2 p}{\partial x^2} + v_p^2\frac{\partial^2 p}{\partial z^2} - 2\eta v_n^2 v_p^2\frac{\partial^4 F}{\partial x^2 z^2} \tag{1}$$

where $F(x,z,t) = \int_o^t \int_0^{t'} p(x,z,\tau)d\tau dt'$, $v_n$ and $v_p$ are the normal moveout (NMO) and compressional velocities, respectively, $p$ is the pressure wavefield, $\varepsilon = \frac{C_{11}-C_{33}}{2C_{33}}$, $\delta = \frac{(C_{13}+C_{44})^2-(C_{33}-C_{44})^2}{2C_{33}(C_{33}-C_{44})}$, $\eta = \frac{\varepsilon-\delta}{1+2\delta}$, $v_n = v_p\sqrt{1+2\delta}$, and the coefficients $C_{11}, C_{33}, C_{44}$, and $C_{13}$ represent equivalent stiffness parameters in the acoustic VTI approximation, introduced to characterize P-wave propagation and anisotropic effects.

It is noted that **Equation 1** is an acoustic wave equation with $\eta = 0$. The space–frequency domain of **Equation 1** can be written as:

$$\omega^2 p + (1+2\eta)v_n^2\frac{\partial^2 p}{\partial x^2} + v_p^2\frac{\partial^2 p}{\partial z^2} - \frac{2\eta v_n^2 v_p^2}{\omega^2}\frac{\partial^4}{\partial x^2\partial z^2}p = 0 \tag{2}$$

**Equation 1** represents a second-order partial differential equation in the depth variable *z*; its equivalent first-order partial differential form can be expressed as follows:

$$\frac{\partial P}{\partial z} = \mathrm{W}P \tag{3}$$

where $P = \begin{bmatrix} p \\ p_z \end{bmatrix}$, $W = \begin{bmatrix} 0 & I \\ H & 0 \end{bmatrix}$, and $H = \frac{\omega^4 + \omega^2(1+2\eta)v_n^2\frac{\partial^2}{\partial x^2}}{\omega^2 v_p^2 - 2\eta v_n^2 v_p^2\frac{\partial^2}{\partial x^2}}$.

**Equation 1** describes the acoustic VTI system in a second-order form, but direct numerical implementation is challenging due to the coupling between temporal and spatial derivatives. Therefore, we converted it into the first-order system shown in **Equation 3**, which enables wavefield extrapolation in the depth direction. This first-order formulation provides two important benefits: (i) it allows explicit eigenvalue decomposition of the system matrix, enabling separation of downward- and upward-propagating components; and (ii) it facilitates the construction of stable numerical schemes suitable for both forward and backward propagation.

The eigendecomposition of $W$ is presented as follows:

$$W = \frac{1}{2}\begin{bmatrix} I & I \\ \Lambda & -\Lambda \end{bmatrix}\begin{bmatrix} \Lambda & \\ & -\Lambda \end{bmatrix}\begin{bmatrix} I & \Lambda^{-1} \\ I & \Lambda^{-1} \end{bmatrix} \tag{4}$$

where $I$ is the identity matrix; $\Lambda$ is the eigenvalue of matrix $W$, with $\Lambda = ik_z$; and $k_z = \sqrt{H}$ is the vertical wavenumber in VTI media.

The mathematical solution to **Equation 4** within VTI media, for the wavefield depth extrapolation, is presented as follows:

$$P(z_j) = e^{W(z_j - z_i)}P(z_i) \tag{5}$$

where $P(z_j) = \begin{bmatrix} \tilde{p}(z_j) \\ \tilde{p}_z(z_j) \end{bmatrix}$ is the pressure and its derivative wavefield at depth $z_j$ levels, and $P(z_i) = \begin{bmatrix} \tilde{p}(z_i) \\ p_z(z_i) \end{bmatrix}$ is the pressure and its derivative wavefield at depth $z_i$ levels.

Furthermore, **Equation 3** can be calculated as:

$$\begin{bmatrix} \tilde{p}(z_j) \\ \tilde{p}(z_j) \end{bmatrix} = \begin{bmatrix} \cos(k_z\Delta z) & \frac{\sin(k_z\Delta z)}{k_z} \\ -k_z\sin(k_z\Delta z) & \cos(k_z\Delta z) \end{bmatrix}\begin{bmatrix} \tilde{p}(z_i) \\ \tilde{p}_z(z_i) \end{bmatrix} \tag{6}$$

where $\Delta z = z_i - z_j$.

A key challenge in using two-way extrapolation in VTI media is the generation of pseudo-S-waves—artificial modes that arise from the acoustic approximation but do not correspond to any physically meaningful propagation. If not properly suppressed, these modes contaminate the extrapolated wavefield and introduce significant migration artifacts. In this study, we integrated an FK-domain filter based on the sign properties of the dispersion relation (Section A1), ensuring that only wave components associated with positive P-wave propagation are retained[32].

Because the proposed scheme performs wavefield depth extrapolation using a two-way wave equation, low-frequency artifacts are produced when using a cross-correlation imaging condition for the source and receiver wavefields. Therefore, the up- and down-going wavefields for source and receiver data need to be separated. Inspired by You *et al.*[42], a new pressure wavefield *q* is defined as:

$$\tilde{q} = \frac{\tilde{p}}{ik_z} \tag{7}$$

Based on **Equation 6**, the novel wavefield depth extrapolation method can be formulated as follows:

$$\begin{bmatrix} \tilde{p}(z_i) \\ \tilde{q}(z_i) \end{bmatrix} = \begin{bmatrix} \cos(k_z \Delta z) & i\sin(k_z \Delta z) \\ i\sin(k_z \Delta z) & \cos(k_z \Delta z) \end{bmatrix} \begin{bmatrix} \tilde{p}(z_i) \\ \tilde{q}(z_i) \end{bmatrix} \tag{8}$$

By using **Equation 7**, the up- and down-going wavefields can be further derived as:

$$\begin{cases} \tilde{p}_d = \frac{1}{2}[\tilde{p} + \tilde{q}] \\ \tilde{p}_u = \frac{1}{2}[\tilde{p} - \tilde{q}] \end{cases} \tag{9}$$

where $\tilde{p}_a$ and $\tilde{p}_d$ are the up- and down-going wavefields, respectively.

During each depth extrapolation, the conventional imaging result can be achieved using wavefield separation based on **Equation 9**, and can be expressed as:

$$Image = \sum_{\omega} \tilde{p}_s^d \cdot \tilde{p}_r^{u*} \tag{10}$$

where $\tilde{p}_s^d$ and $\tilde{p}_r^u$ represent the down-going source wavefield and up-going receiver wavefield, respectively, and * denotes the conjugate operator.

Based on the above algorithm description, the specific implementation is shown in Table 1.

## 3. Results

To comprehensively evaluate the performance of the proposed VTI two-way extrapolation method, we designed numerical experiments spanning a range of geological complexities—from layered media to strongly faulted structures and real field data situations. The goal of these tests is not only to validate the numerical stability of the method but also to demonstrate its advantages in resolving phase dispersion, maintaining reflector continuity, and preserving amplitude relative to conventional acoustic migration.

### 3.1. Three-layer model

A three-layer synthetic model, illustrated in Figure 1, was constructed with spatial dimensions of 3,000 m in depth (z) and 6,000 m in width (x). The grid sampling intervals were set to 10.0 m vertically and 20.0 m horizontally. A Ricker wavelet centered at 20.0 Hz served as the source function. Seismic data acquisition spanned 2.0 s with a temporal sampling rate of 0.001 s. Receivers were uniformly placed along the surface from 0 to 6,000 m at 20.0 m intervals. Additionally, 150 shot points were arranged with a spacing of 40.0 m.

The simulated shot gathers for both acoustic and VTI media are shown in Figure 2. At a representative location (x = 1,000 m), the waveform response in the VTI medium exhibited pronounced phase dispersion

**Table 1. Two-way vertically transversely isotropic wavefield depth extrapolation and imaging**

| Category | Description |
|---|---|
| Input | Acoustic VTI model parameters $v_{nmo}(\mathrm{x},\mathrm{z})$, $v_p(\mathrm{x},\mathrm{z})$, anisotropic parameters $\varepsilon(\mathrm{x},\mathrm{z})$, $\delta(\mathrm{x},\mathrm{z})$;<br>Source wavelet $s(t)$;<br>Recorded seismic data $\mathrm{d}(\mathrm{x}_r,t)$;<br>Depth step $\Delta z$;<br>Frequency sampling $\omega$. |
| Steps | 1) Initialization: Transform the source wavelet and recorded seismic data from the time domain to the frequency domain. Initialize the source and receiver wavefields at the surface depth $z = 0$;<br>Depth extrapolation: For each depth step, extrapolate both source and receiver wavefields downward using the two-way wavefield extrapolation operator (**Equation 6**);<br>Pseudo-S-wave and evanescent-wave suppression: Apply a frequency–wavenumber-domain filter to the extrapolated wavefields to remove pseudo-S-wave and evanescent components based on the dispersion relation criteria (Section A1);<br>Wavefield transformation: Compute the auxiliary pressure wavefield using **Equation 7** to facilitate stable separation of wavefield components;<br>Up- and down-going wavefield separation: Decompose the source and receiver wavefields into down-going and up-going components using **Equation 9**;<br>Imaging condition: Apply the cross-correlation imaging condition by correlating the down-going source wavefield with the up-going receiver wavefield at each depth level (**Equation 10**);<br>Image accumulation: Sum the imaging contributions over all frequencies and shots to obtain the final migrated image $I(x,z)$. |
| Output | Depth-domain migrated image $I(x,z)$. |

relative to its acoustic counterpart, as seen in Figure 2C. An enlarged view of the region highlighted by the red-dashed rectangle in Figure 2C is shown in Figure 2D to more clearly illustrate the phase discrepancies. Using these VTI data, both a conventional acoustic migration and the proposed VTI-based migration method were performed. Figure 3 presents the resulting images. The conventional acoustic approach produced a structurally correct image but exhibited significant phase-dispersion artifacts in imaging results. In contrast, the VTI scheme yielded a more coherent image with minimal dispersion artifacts.

In the three-layer example, the model was intentionally kept simple to clearly observe the influence of VTI anisotropy. By comparing the acoustic and VTI shot gathers, we observed systematic phase discrepancies originating from directional velocity differences introduced by ε and δ. These discrepancies propagated directly into the migration results: the acoustic scheme yielded smeared reflectors, while the VTI scheme reconstructed sharper interfaces. This example highlights the fundamental necessity of anisotropy correction even in structurally simple environments.

### 3.2. Hess model

An anisotropic Hess model was used for imaging comparison, as shown in Figure 4, with a vertical extent of 2,500 m and a horizontal span of 10,000 m. The spatial sampling intervals were set to 10.0 m in the vertical direction and 20.0 m horizontally. The source was represented by a 20.0 Hz Ricker wavelet. Data acquisition spanned 3.0 s with samples recorded every 0.001 s. Receivers were positioned uniformly along the surface from 0 to 10,000 m at 20.0 m intervals, while 250 shot points were placed every 40.0 m.

Modeled shot gathers for acoustic and VTI media are displayed in Figure 5. At x = 6,000 m, the VTI simulation revealed significant phase dispersion compared to the acoustic case, as shown in Figure 5C. Figure 5D zooms into the region outlined by the red-dashed rectangle in Figure 5C, providing a clearer view of the phase variation between the two media. To assess migration accuracy, both the proposed VTI method and the conventional acoustic approach were applied to data generated from VTI models. As illustrated in Figure 6, the VTI migration produces clearer, more geologically consistent images with fewer artifacts and improved structural continuity, as indicated by the red-dashed ellipse. For the Hess VTI model, which contained strong lateral velocity gradients and dipping structures, the advantage of the proposed method became more pronounced. Acoustic migration visibly distorted reflectors beneath high-velocity regions, particularly in areas with strong anisotropy. The proposed VTI method significantly improved the continuity of deep reflectors and reduced migration noise in the model's flanks, as highlighted by the red arrows in Figure 6. The improvement in imaging accuracy was not only quantitative but also qualitative in regions that were typically difficult for one-way acoustic methods to handle.

### 3.3. Marmousi model

The anisotropic Marmousi model, shown in Figure 7, spanned 2,000 m in depth and 8,000 m in width, with vertical and horizontal sampling intervals of 10.0 m and 20.0 m, respectively. A Ricker wavelet centered at 20.0 Hz was used to excite the model. Seismic recordings extended

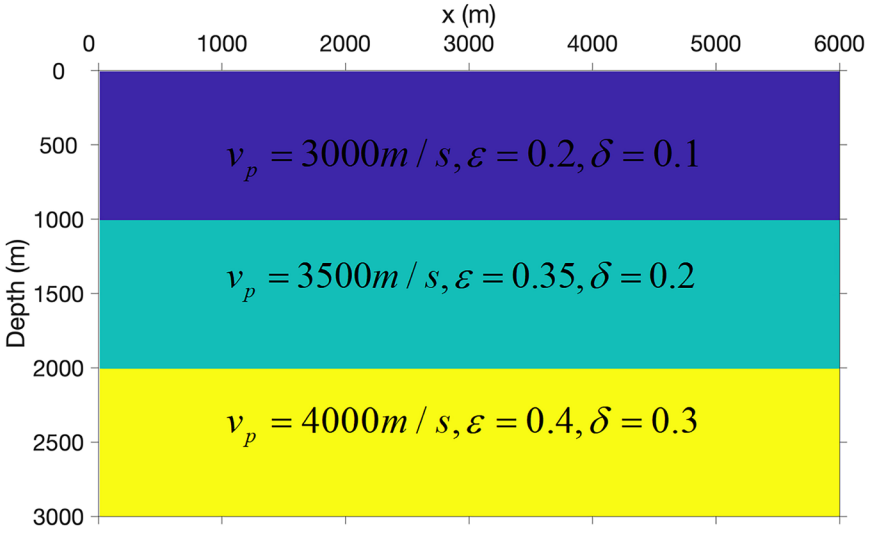



**Figure 1.** Anisotropy parameters of the three-layer model

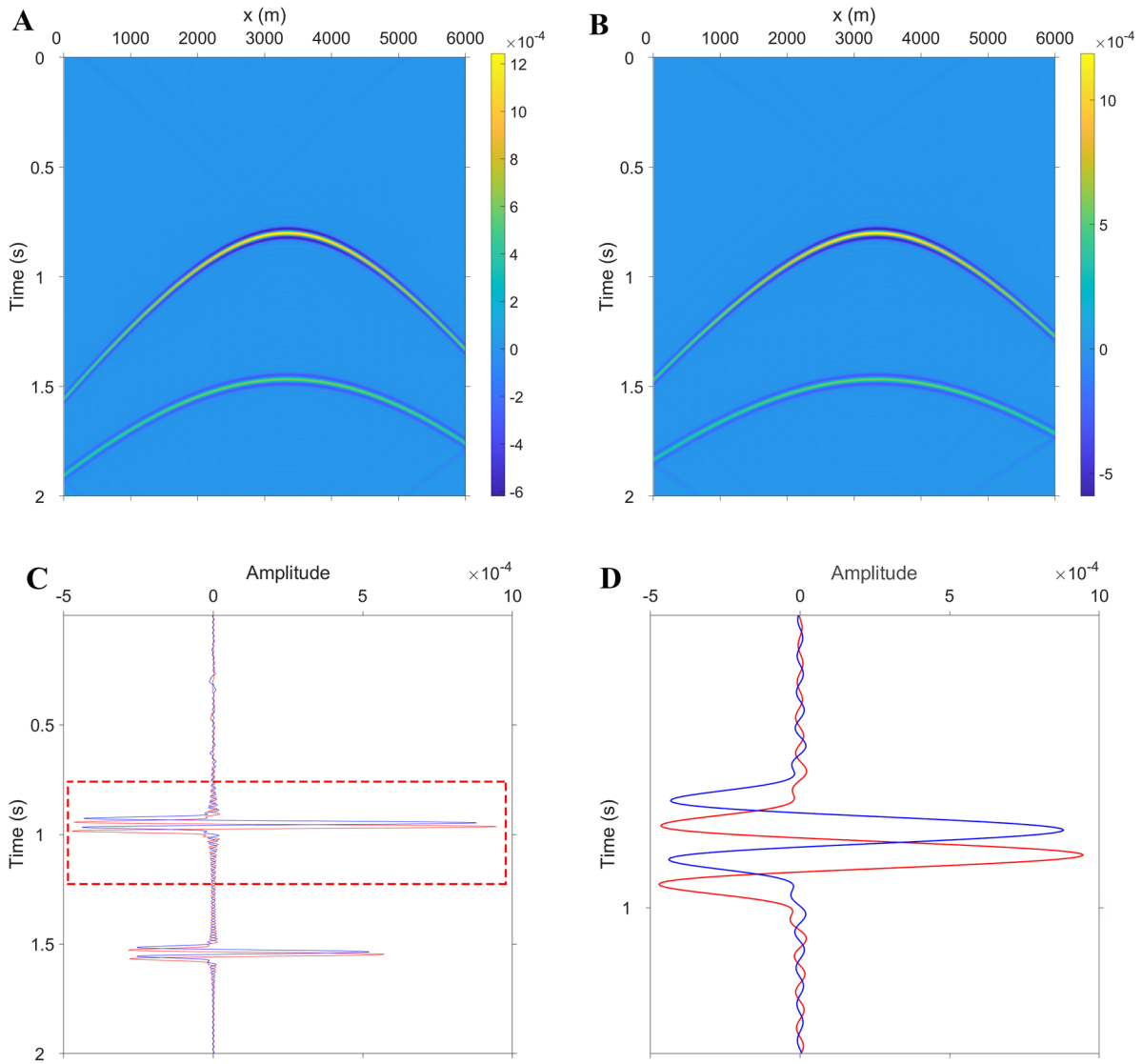



**Figure 2.** Shot gather obtained in (A) acoustic media and (B) vertically transversely isotropic (VTI) media. (C) Waveform comparison. (D) Enlarged view of the area highlighted by the red-dashed rectangle in (C). The red line denotes waveforms calculated in acoustic media, while the blue line presents waveforms calculated in VTI media.

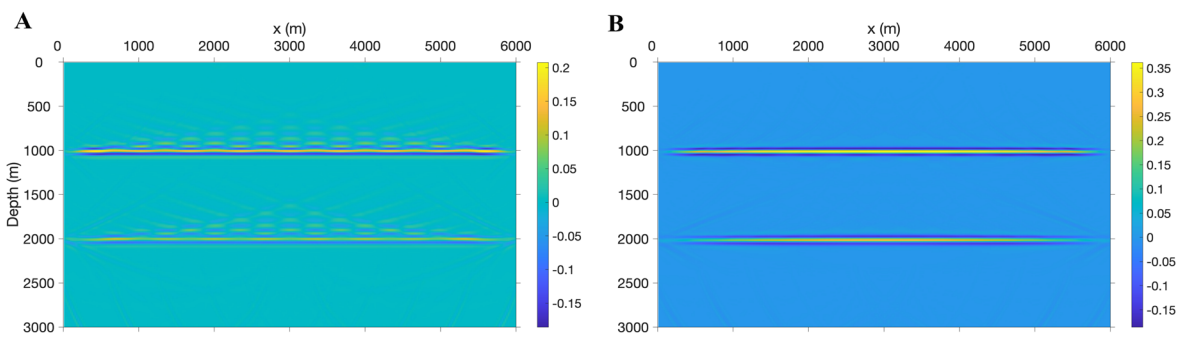



**Figure 3.** Imaging results of shot gathers obtained from (A) acoustic media and (B) vertically transversely isotropic media

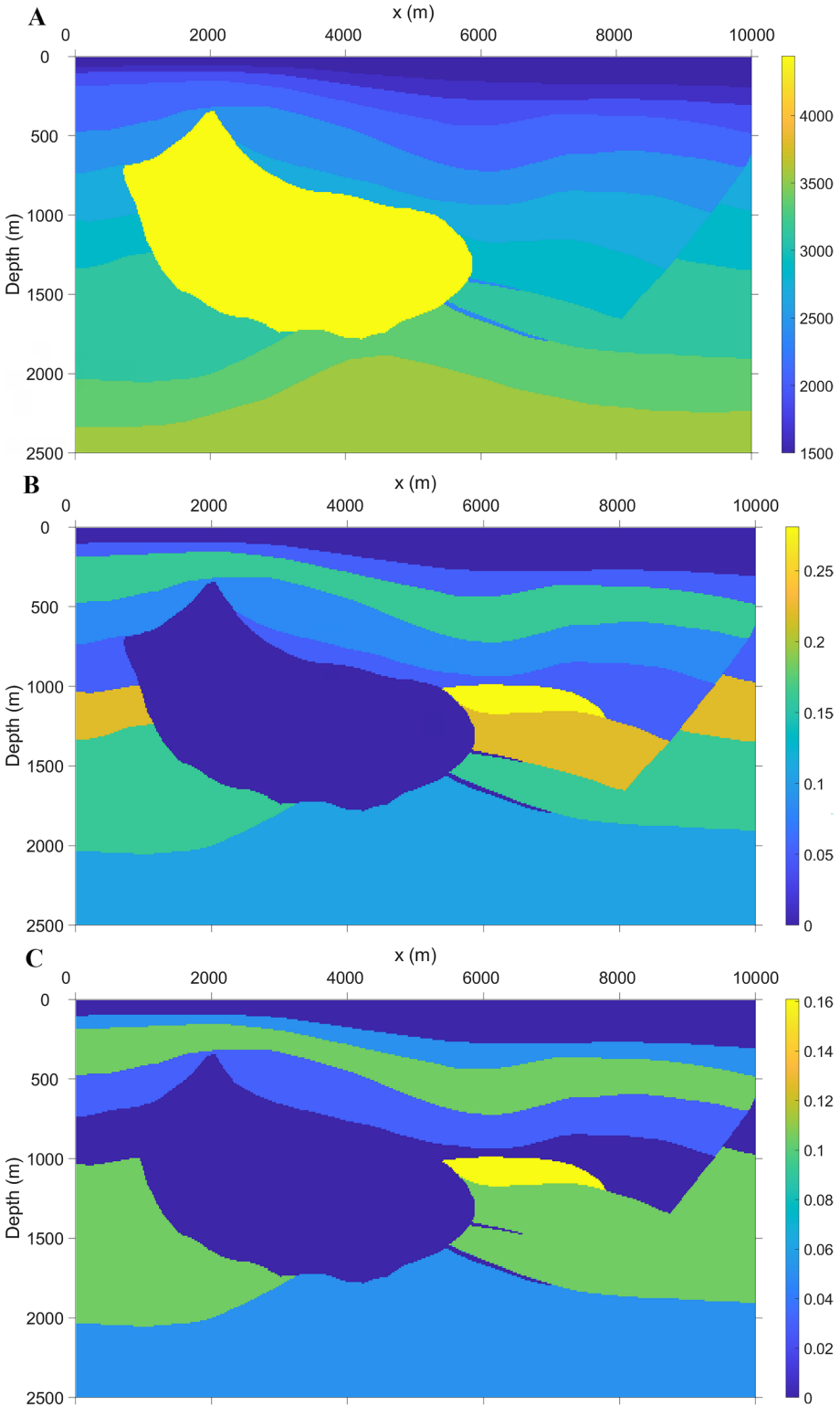



**Figure 4.** Hess model parameters: (A) P-wave velocity, (B) anisotropy parameter $\varepsilon$, and (C) anisotropy parameter $\delta$

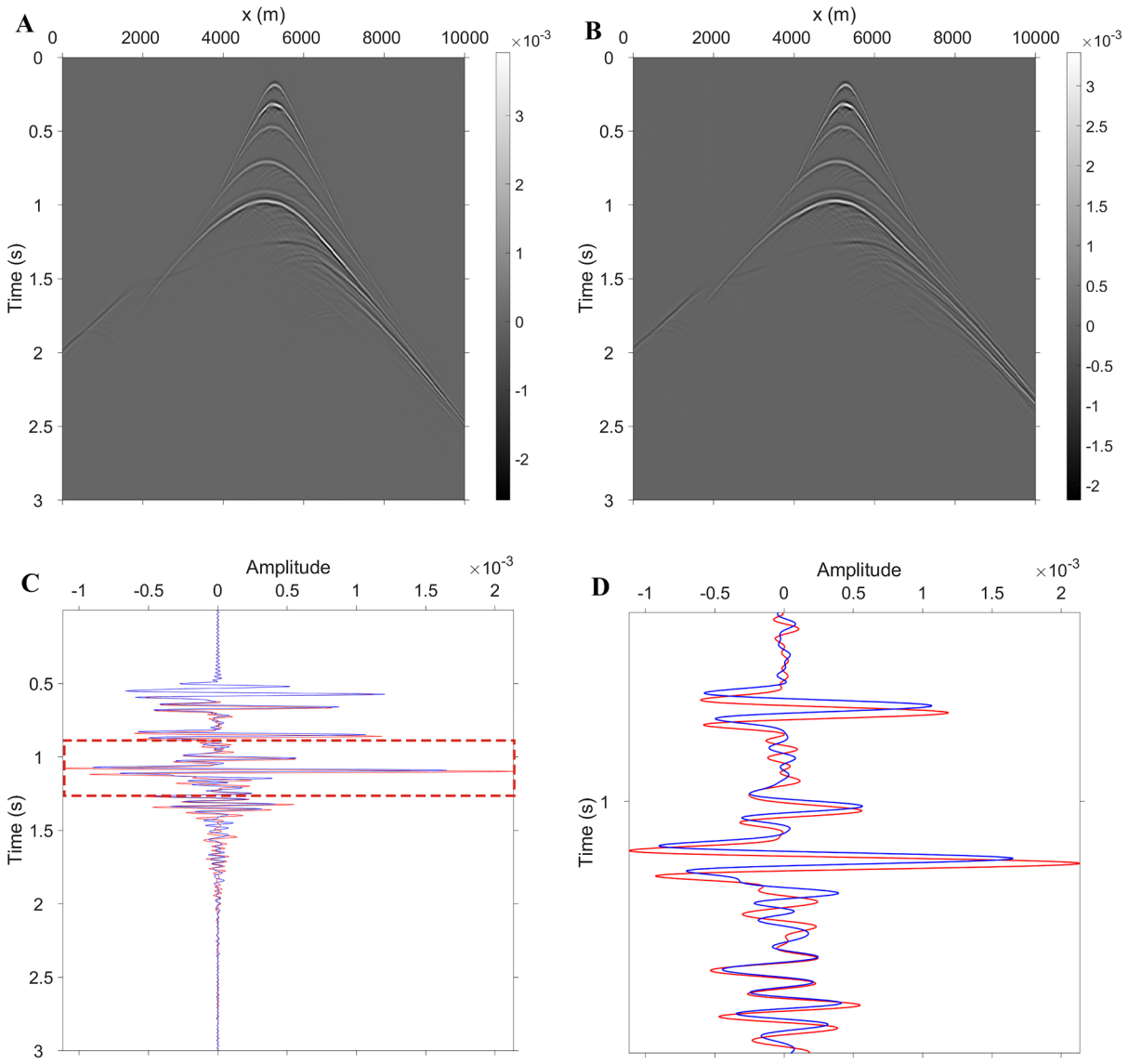



**Figure 5.** Shot gathers obtained from (A) acoustic media and (B) vertically transversely isotropic (VTI) media. (C) Waveform comparison. (D) Enlarged view of the area highlighted by the red-dashed rectangle in (C). The red line denotes waveforms calculated in acoustic media, while the blue line presents waveforms calculated in VTI media.

to 3.0 s with a temporal resolution of 0.001 s. The receiver array covered the entire model top from 0 to 8,000 m with a 20.0 m spacing, and 200 shot points were distributed at 40.0 m spacing.

Figure 8 illustrates the shot gathers for both acoustic and VTI conditions. At x = 5,000 m, the VTI medium again displayed clear phase dispersion, in contrast to the acoustic case, as depicted in Figure 8C. To emphasize the phase differences, an enlarged view of the highlighted region in Figure 8C is provided in Figure 8D, allowing for a more detailed comparison of the waveform behavior between the two media.

For migration, both the proposed VTI-based method and the conventional acoustic strategy were employed. The resulting images in Figure 9 show that the acoustic scheme failed to resolve complex structures between the three major faults. The Marmousi VTI model further tested the method's ability to recover complex fault geometries and thin-bed structures. The acoustic migration failed to reconstruct reflector terminations between major faults, whereas the VTI method recovered them with greater clarity and consistency, as indicated by the red-dashed

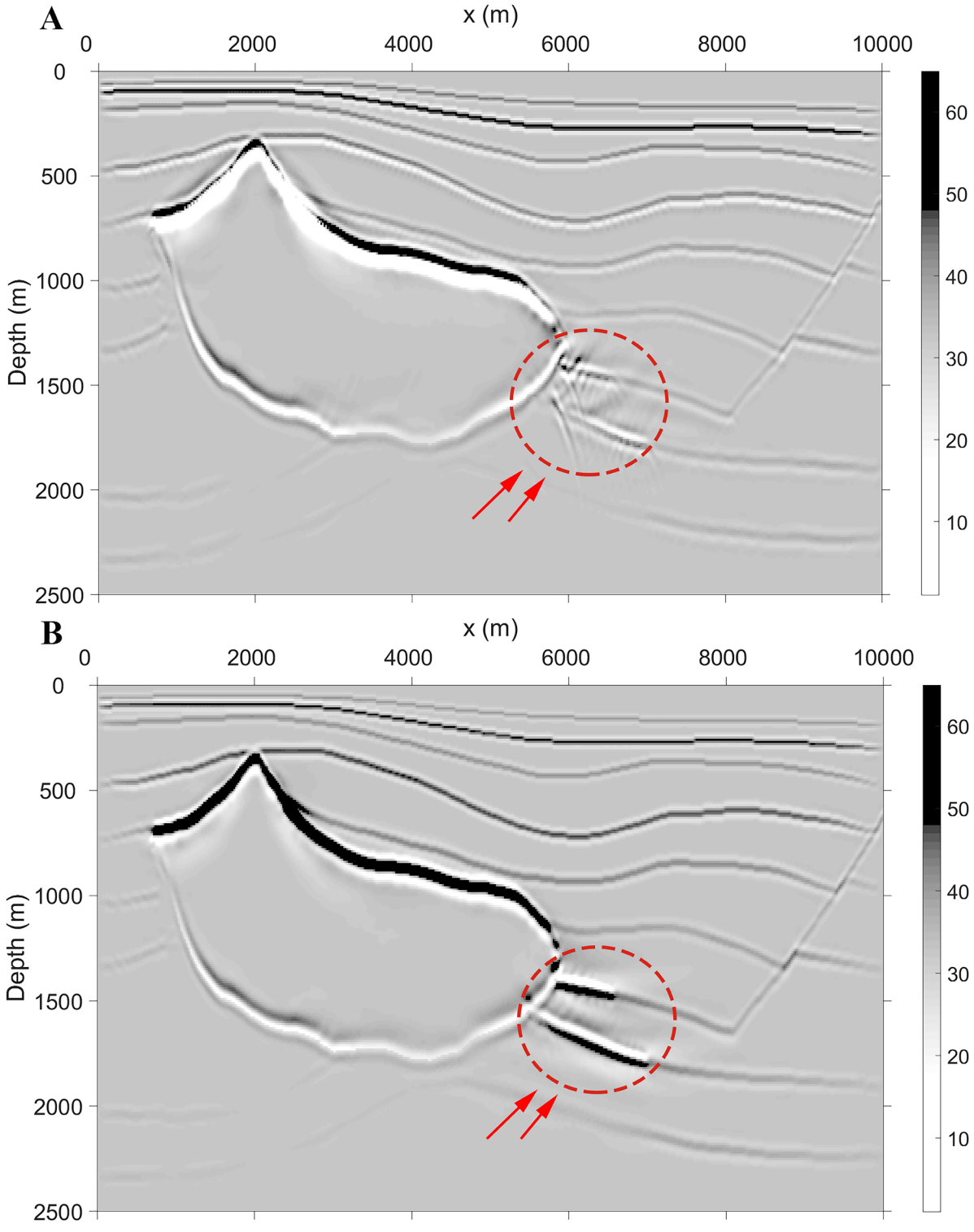



**Figure 6.** Imaging results of shot gather obtained from (A) acoustic media and (B) vertically transversely isotropic media. The red-dashed ellipse and arrows illustrate the imaging differences.

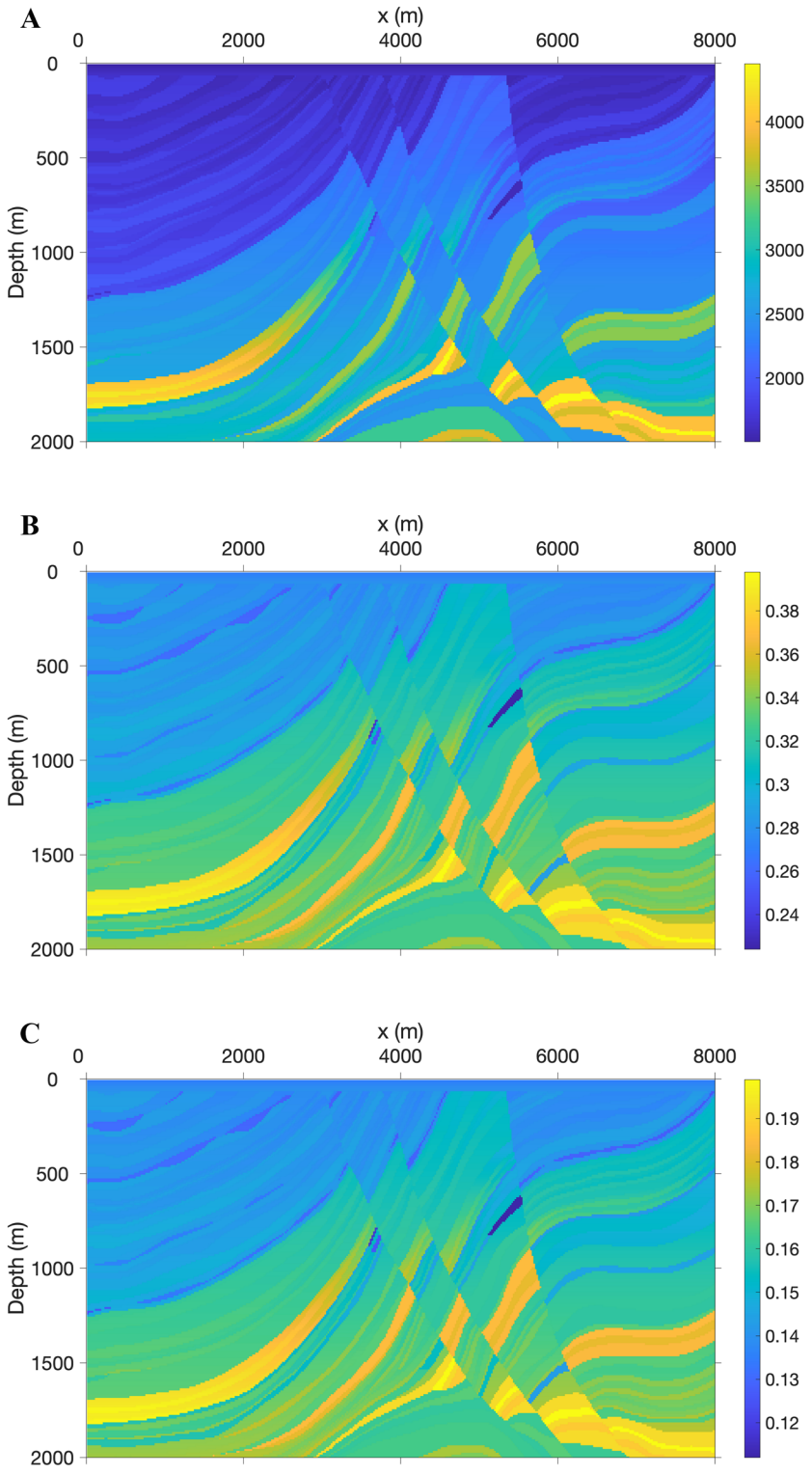



**Figure 7.** Marmousi model parameters: (A) P-wave velocity, (B) anisotropy parameter $\varepsilon$ , and (C) anisotropy parameter $\delta$

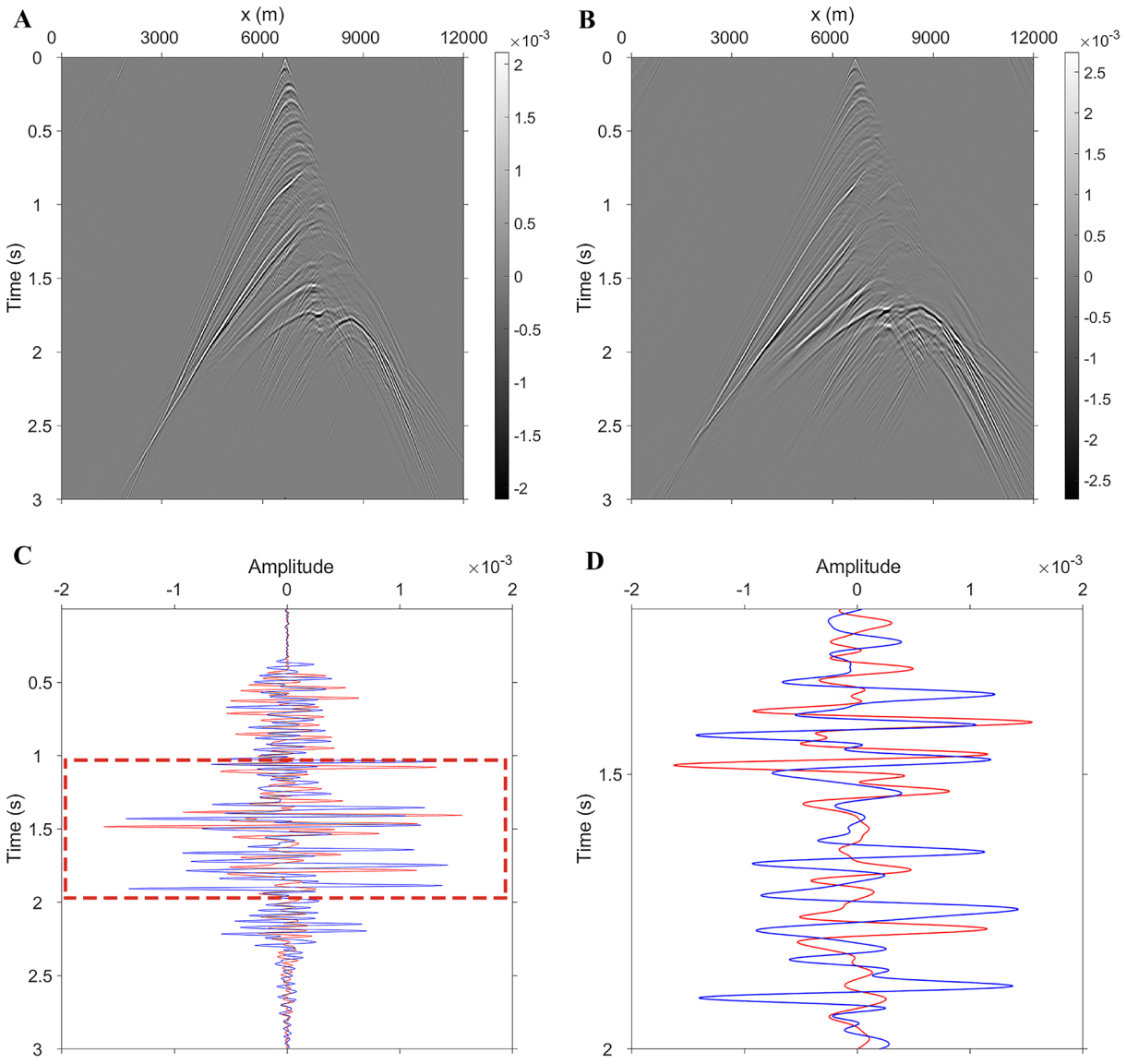



**Figure 8.** Shot gather obtained from (A) acoustic media and (B) vertically transversely isotropic (VTI) media. (C) Waveform comparison. (D) Enlarged view of the area highlighted by the red-dashed rectangle in (C). The red line denotes waveforms calculated in acoustic media, while the blue line presents waveforms calculated in VTI media.

ellipse. These results indicate that the proposed two-way extrapolation can maintain accurate phase propagation even under strong lateral heterogeneity and anisotropic variations. To further highlight local differences, Figure 10 presents zoomed-in views of representative regions. Figure 10A,B correspond to acoustic and VTI media, respectively, demonstrating that the VTI method better preserved reflector continuity and phase characteristics at the small scale.

### 3.4. Real data application

To evaluate the effectiveness of the proposed method on real data, we applied our imaging workflow to a land seismic dataset obtained from previous studies.[47,48] The shot gathers used in this study are shown in Figure 11. As observed from the gathers, the continuity of reflection events was relatively poor, particularly for deeper reflections, which is typical for land data acquired in geologically complex areas.

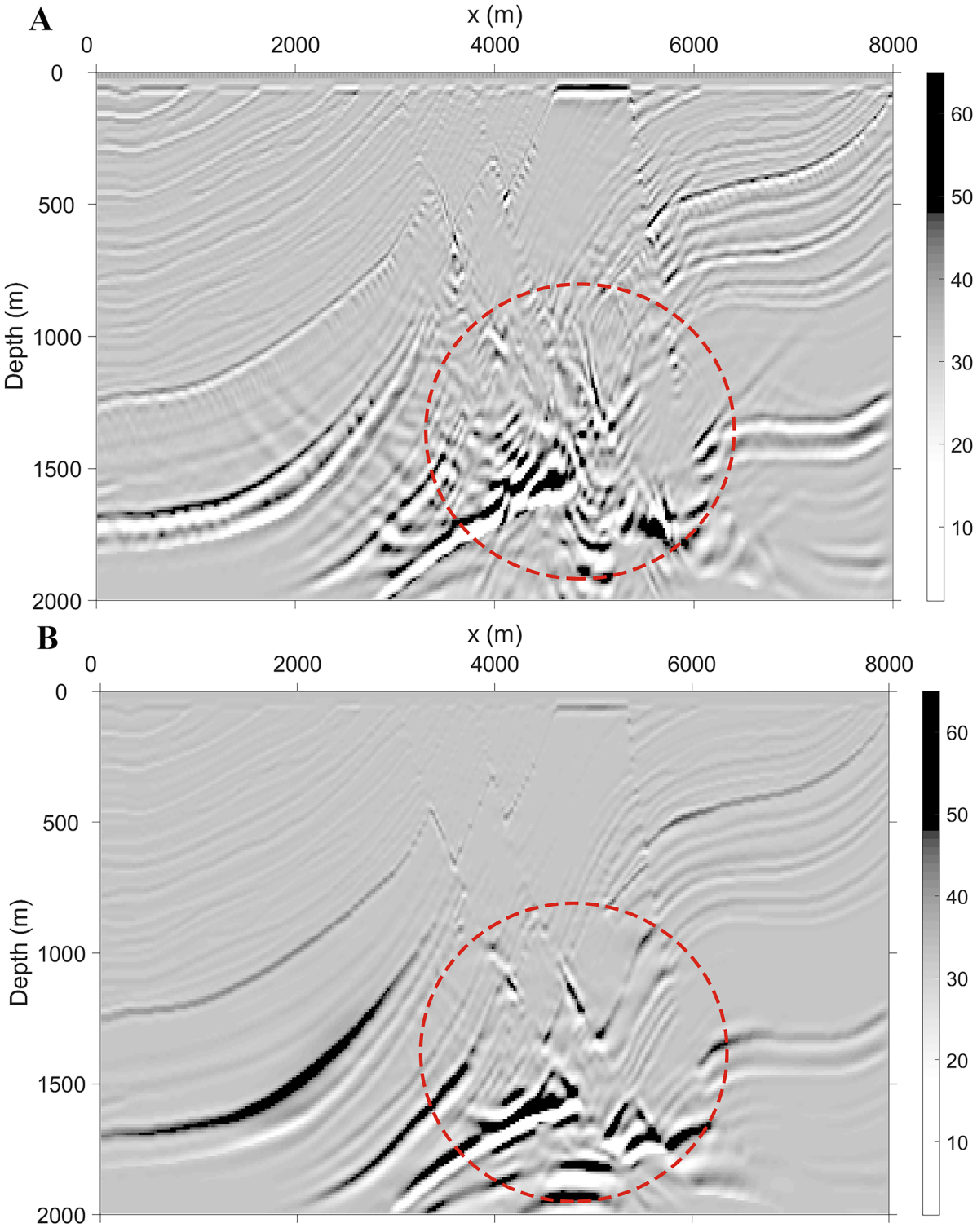



**Figure 9.** Imaging results of shot gather obtained from (A) acoustic media and (B) vertically transversely isotropic media. The red-dashed ellipse shows the imaging differences.

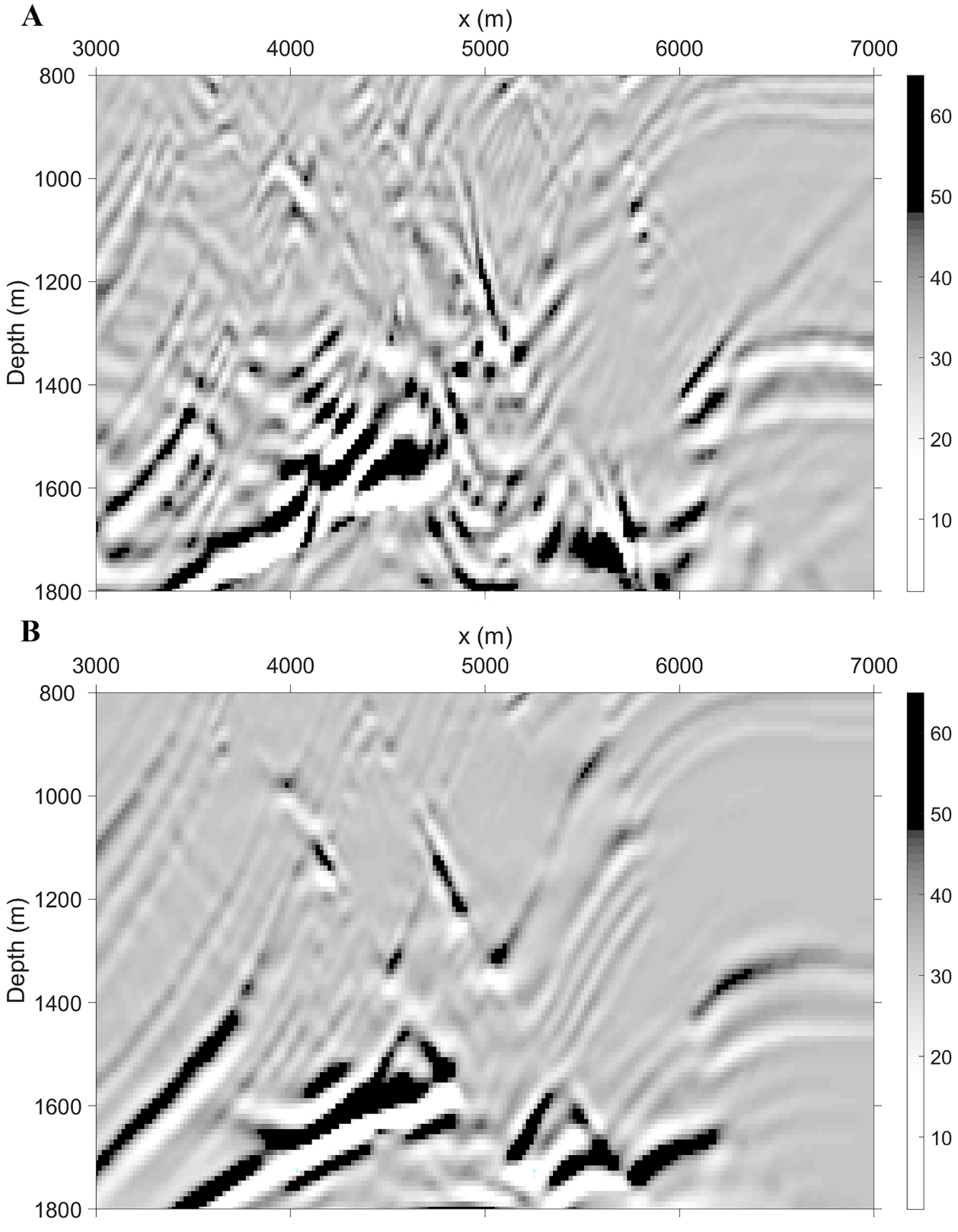



**Figure 10.** Local imaging results from (A) acoustic media and (B) vertically transversely isotropic media

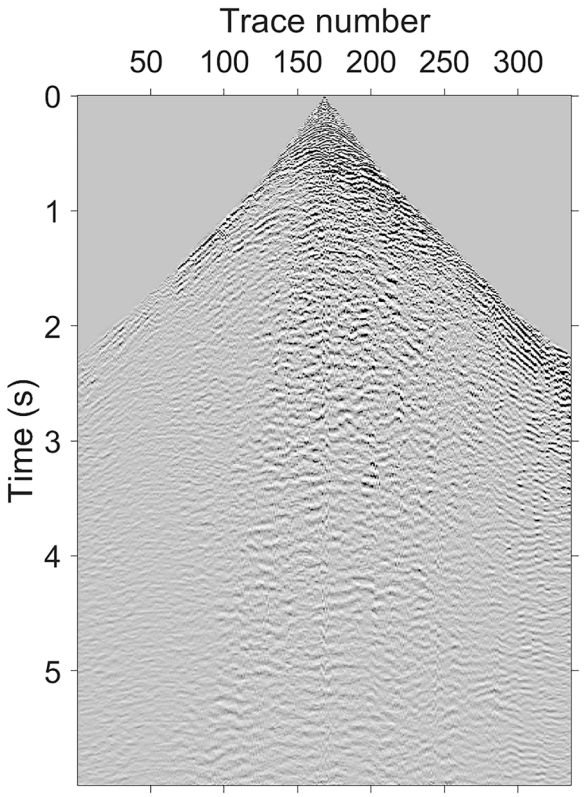



**Figure 11.** Real land seismic shot gathers

A depth-domain velocity model, together with the corresponding epsilon and delta fields, was constructed using a conventional VTI traveltime inversion scheme (Figure 12).[49,50] Using this model, we performed depth imaging of the shot gathers with two different two-way wave-equation-based migration approaches: one assuming an acoustic (isotropic) medium and the other incorporating VTI anisotropy. The resulting migrated images are presented in Figure 13.

From the perspective of structural imaging, the two-way wave-equation-based migration method for VTI media developed in this study produced a noticeably more coherent and geologically consistent image, as indicated by the red-dashed ellipse. Compared with the acoustic migration result, the VTI migration enhanced the continuity of key structural reflectors, as highlighted by the red arrows in Figure 13. To further illustrate these differences, Figure 14 shows zoomed-in views of the two regions indicated by the red ellipses in Figure 13, where the VTI results exhibited clearer reflector terminations and improved local continuity compared with the acoustic counterparts. The imaging performance on this real land dataset demonstrates the practical advantages of the proposed VTI two-way propagator in resolving subsurface structures and improving reflector continuity.

## 4. Discussion

A key novelty of our work lies in the explicit construction of a two-way VTI propagator based on the first-order system derived from the acoustic VTI wave equation. Traditional acoustic depth-migration methods assume isotropy during wavefield extrapolation, leading to phase inaccuracies when applied to anisotropic models. Moreover, the coupling between P- and pseudo-S-wave components becomes particularly problematic when depth extrapolation is performed using two-way equations, because the imaging condition tends to amplify low-wavenumber artifacts. By integrating an FK-domain filtering strategy into the extrapolation process, we successfully eliminated pseudo-S-wave and evanescent-wave energy while retaining the physically meaningful P-wave response. This step is crucial for obtaining stable, interpretable imaging results. It represents a key methodological advancement over classical two-way extrapolation schemes.

Despite these improvements, some limitations remain. First, the current formulation uses an acoustic VTI approximation and therefore does not capture any real shear-wave physics and therefore neglects elastic mode

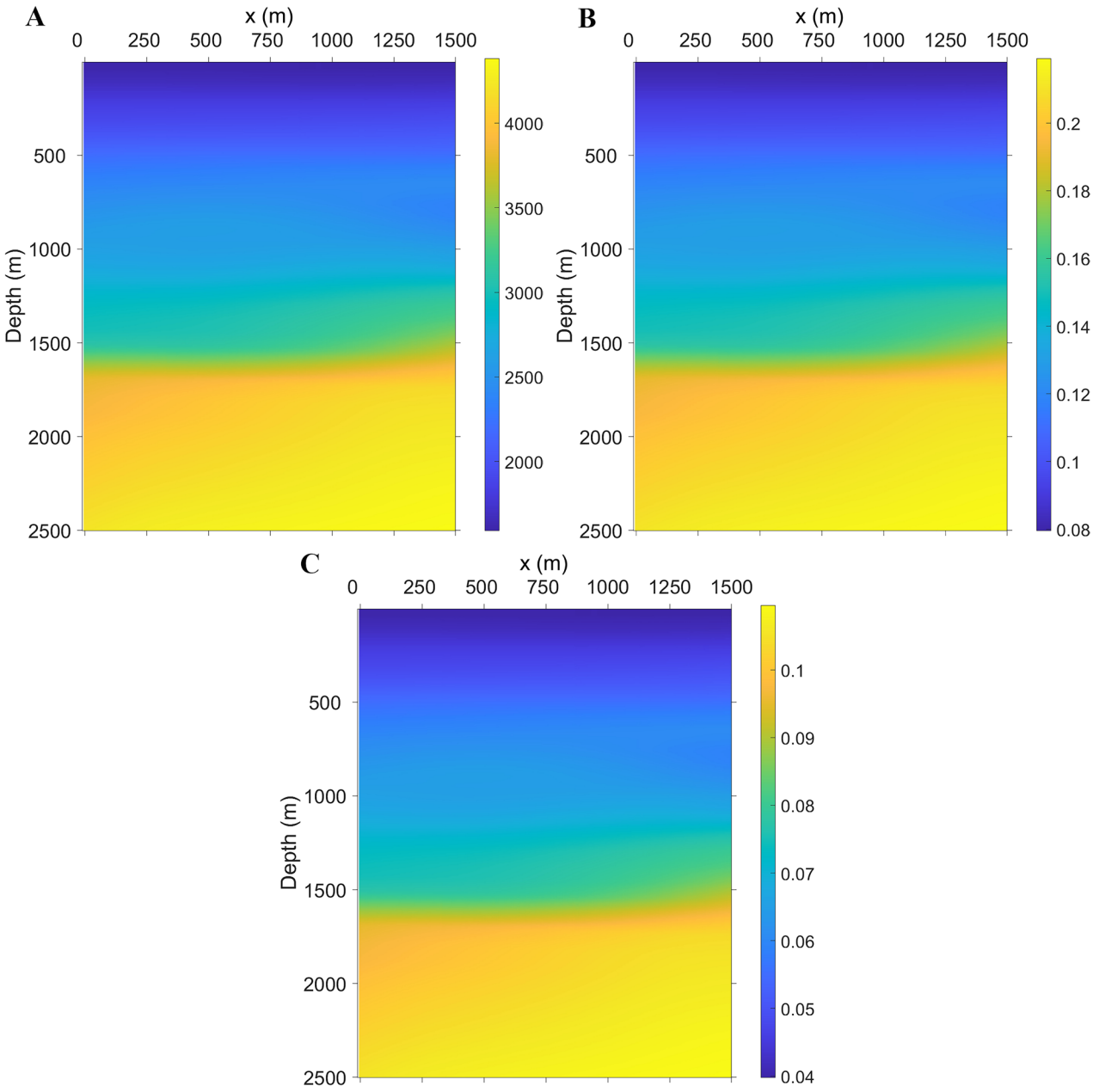



**Figure 12.** Real data parameters: (A) P-wave velocity, (B) anisotropy parameter $\varepsilon$, and (C) anisotropy parameter $\delta$

conversions and qS-related responses. While the pseudo-S suppression strategy works well for P-wave imaging, this approximation may limit the method's applicability in areas where elastic effects are non-negligible, and incorporating elastic effects or qSV-sensitive reflections may require extending the method to a fully elastic two-way formulation. Second, the method's accuracy depends heavily on the quality of the anisotropic parameters (ε and δ), which are often difficult to estimate reliably in practical applications. Errors in anisotropic model building may propagate into the extrapolated wavefields, degrading imaging performance and potentially leading to residual artifacts and reduced image focusing. Therefore, tighter integration with anisotropic velocity/parameter inversion—such as reflection-based waveform inversion or PINN-based parameter estimation—is essential to improve the method's robustness in field-data applications.

Future research may focus on several aspects. One direction is to generalize the proposed two-way propagator

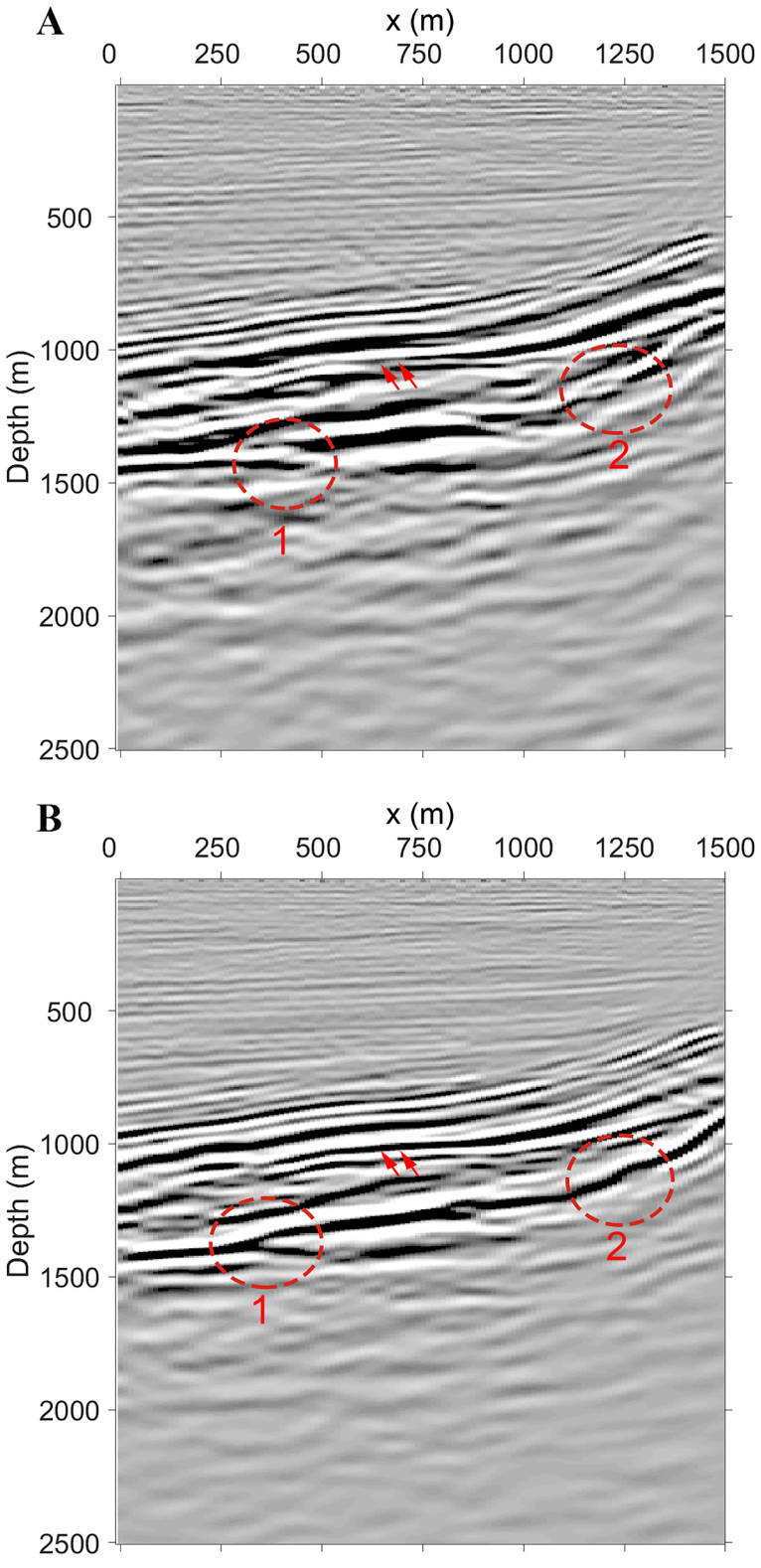



**Figure 13.** Imaging results of shot gather obtained from (A) acoustic media and (B) vertically transversely isotropic media. The red-dashed ellipse and arrows indicate the imaging differences.

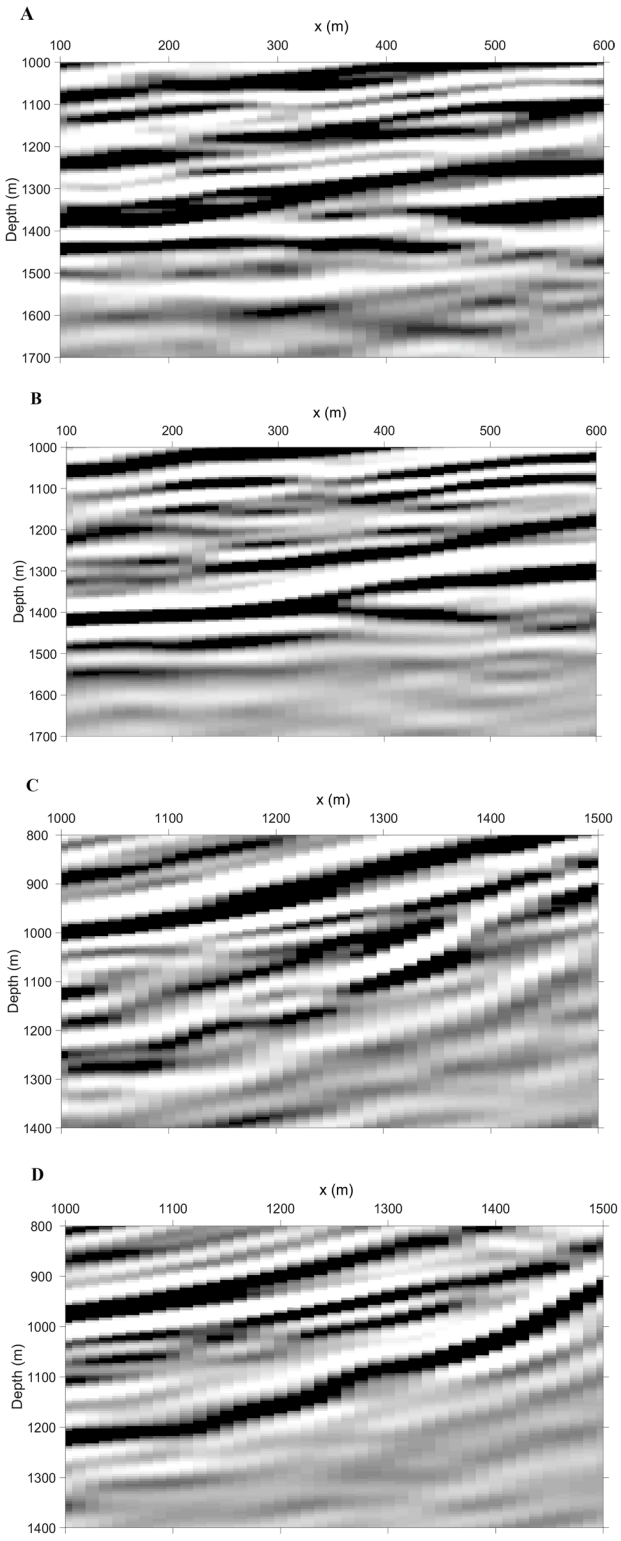



**Figure 14.** Local imaging results for the regions indicated by red ellipses in Figure 13. (A) Red ellipse 1 in Figure 13A. (B) Red ellipse 1 in Figure 13B. (C) Red ellipse 2 in Figure 13A. (D) Red ellipse 2 in Figure 13B.

to tilted-transverse isotropy media, which are often required for imaging complex geologic structures, such as sub-salt regions and thrust belts. Another promising extension is the development of joint velocity and anisotropy estimation workflows based on the proposed two-way operator, potentially using gradient-based inversion or machine-learning-enhanced optimization. Furthermore, combining the proposed propagator with least-squares migration could enhance amplitude fidelity and improve the recovery of weak reflections. Finally, GPU-parallelized implementations may significantly accelerate the computational efficiency of the two-way VTI extrapolation operator, making it more suitable for large-scale industrial applications.

## 5. Conclusion

To address the challenge of accurate seismic imaging in anisotropic media, this study proposed a novel acoustic two-way wave-equation wavefield depth extrapolation method specifically designed for imaging in VTI media. Extensive simulations verified that the proposed VTI imaging strategy, based on the two-way wave equation, consistently outperformed the traditional acoustic approach in processing data from anisotropic media. Across a range of geological scenarios—from a simple three-layer model to the velocity-gradient Hess model and the structurally complex Marmousi model—the conventional acoustic method produced images with notable phase distortion and structural inaccuracies because it could not accommodate anisotropic effects. In contrast, the VTI scheme yielded significantly improved imaging results, effectively suppressing phase dispersion and enhancing structural continuity, particularly in geologically complex regions. These findings underscore the limitations of conventional acoustic assumptions for representing the anisotropic nature of real Earth media and highlight the need for VTI-based methods to achieve high-fidelity seismic imaging.

## Acknowledgments

None.

## Funding

This work was supported by the open research project of the National Engineering Research Center for Oil and Gas Exploration Computer Software (DFWT-ZYRJ-2025-JS-56) and Sichuan Science and Technology Program (2026NSFSC0234, 2024NSFSC0810).

## Conflict of interest

Jiachun You is a Guest Editor of this special issue, but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Author contributions

*Conceptualization:* Shanzheng Hu
*Formal analysis:* Pengyuan Sun
*Investigation:* Jiachun You
*Methodology:* Jiachun You
*Writing–original draft:* Shuqin Li
*Writing–review & editing:* Feng Hu, Min Guan

## Availability of data

Data will be made available upon request to the corresponding author.

## References

1. Alkhalifah T. Acoustic approximations for processing in transversely isotropic media. *Geophysics*. 1998;63(2):623-631.

   doi: 10.1190/1.1444361

2. Alkhalifah T. An acoustic wave equation for anisotropic media. *Geophysics*. 2000;65(4):1239-1250.

   doi: 10.1190/1.1444815

3. Li ZC, Qu YM. Research progress on seismic imaging technology. *Pet Sci*. 2022;19(1):128-146.

   doi: 10.1016/j.petsci.2022.01.015

4. Duveneck E, Milcik P, Bakker PM, Perkins C. Acoustic VTI wave equations and their application for anisotropic reverse-time migration. In: *SEG Technical Program Expanded Abstracts 2008*; Society of Exploration Geophysicists; 2008:2186-2190.

   doi: 10.1190/1.3059320

5. Bube KP, Nemeth T, Stefani JP, *et al*. On the instability in second-order systems for acoustic VTI and TTI media. *Geophysics*. 2012;77(5):T171-T186.

   doi: 10.1190/geo2011-0250.1

6. Zhou H, Zhang G, Bloor R. An Anisotropic Acoustic Wave Equation for VTI Media. In: *Proceedings of the 68th EAGE Conference and Exhibition Incorporating SPE EUROPEC 2006*; June 12-15, 2006; Vienna, Austria. European Association of Geoscientists & Engineers; 2006:cp-2-00318.

   doi: 10.3997/2214-4609.201402310

7. Yan J, Sava P. Elastic wave-mode separation for VTI media. *Geophysics*. 2009;74(5):WB19-WB32.

   doi: 10.1190/1.3184014

8. Zhang Q, McMechan GA. 2D and 3D elastic wavefield vector decomposition in the wavenumber domain for VTI media. *Geophysics*. 2010;75(3):D13-D26.

   doi: 10.1190/1.3431045

9. Yao G, Fang X, Zheng Q, *et al*. Pseudo-Helmholtz decomposition for an elastic VTI wavefield based on wavefront phase direction. *Geophysics*. 2024;89(3):T151-T162.

   doi: 10.1190/geo2023-0574.1

10. Sun S, Mao W, Zhang Q, *et al*. Reverse-time migration for pure qP-wave based on elliptical decomposition vector equation. *IEEE Trans Geosci Remote Sens*. 2024;62:1-12.

    doi: 10.1109/TGRS.2024.3406763

11. Shan T, Yang J, Huang J, Wang W, Qin S. Least-Squares Gaussian Beam Migration in the VTI Media With a Cauchy Constrain. *IEEE Trans Geosci Remote Sens*. 2025;63:1-9.

    doi: 10.1109/tgrs.2025.3554996

12. Li A, Li C, Lai F, Liao J, Liu T. An Acoustic Wave-Equation Depth Migration Method Using Generalized Two-Way Phase Shift Operator for Anisotropic Media. *IEEE Geosci Remote Sens Lett*. 2025;22:1-5.

    doi: 10.1109/lgrs.2025.3549919

13. Qin S, Yang J, Wang P, Huang J, Qin N. Seismic imaging of pure quasi-P-wave in the VTI media by using the optical flow to calculate phase-velocity direction. *Geophys J Int*. 2025;242(3):ggaf270.

    doi: 10.1093/gji/ggaf270

14. Zhang K, Guo LJ, Li ZC, *et al*. Reverse time migration of VTI media in undulating surface based on the first-order velocity-stress equation. *Geophys Prospect Petr*. 2026;65(1):77-87. [In Chinese].

    doi: 10.12431/issn.1000-1441.2024.0135

15. Fomel S, Ying L, Song X. Seismic wave extrapolation using lowrank symbol approximation. *Geophys Prospect*. 2013;61(3):526-536.

    doi: 10.1111/j.1365-2478.2012.01064.x

16. Sun H, Sun J. A VTI medium prestack migration method based on the De Wolf approximation. *Comput Geosci*. 2025;196:105835.

    doi: 10.1016/j.cageo.2024.105835.

17. Mao Q, Huang JP, Mu XR, *et al*. Accurate simulations of pure-viscoacoustic wave propagation in tilted transversely isotropic media. *Pet Sci*. 2024;21(2):866-884.

    doi: 10.1016/j.petsci.2023.11.005

18. Liang K, Cao D, Sun S, *et al*. Decoupled wave equation and forward modeling of qP wave in VTI media with the new acoustic approximation. *Geophysics*. 2023;88(1):WA335-WA344.

    doi: 10.1190/geo2022-0292.1

19. Bai T, Zhu T, Tsvankin I. Attenuation compensation for time-reversal imaging in VTI media. *Geophysics*. 2019;84(4):C205-C216.

    doi: 10.1190/geo2018-0532.1

20. Qu Y, Zhu J, Chen Z, *et al*. Q-compensated least-squares reverse time migration with velocity-anisotropy correction based on the first-order velocity-pressure equations. *Geophysics*. 2022;87(6):S335-S350.

    doi: 10.1190/geo2021-0689.1

21. Wang N, Xing G, Zhu T, *et al*. Propagating seismic waves in VTI attenuating media using fractional viscoelastic wave equation. *J Geophys Res Solid Earth*. 2022;127(4):e2021JB023280.

    doi: 10.1029/2021jb023280.

22. Wu S, Wang T, Cheng J. Second-order optimization for multiparameter reflection waveform inversion in acoustic VTI media. *Geophys J Int*. 2024;236(1):249-269.

    doi: 10.1093/gji/ggad406

23. Singh S, Tsvankin I, Zabihi Naeini E. Full-waveform inversion with borehole constraints for elastic VTI media. *Geophysics*. 2020;85(6):R553-R563.

    doi: 10.1190/geo2019-0816.1

24. Lang K, Yin X, Zong Z, *et al*. Anisotropic nonlinear inversion based on a novel PP wave reflection coefficient for VTI media. *IEEE Trans Geosci Remote Sens*. 2023;61:1-13.

    doi: 10.1109/TGRS.2023.3295800

25. Wang B, Zhang F, Dai FC, *et al*. Study on seismic inversion method of SH-SH wave in VTI media. *Chin J Geophys*. 2023;66(5):2112-2122. [In Chinese].

    doi: 10.6038/cjg2022P0750

26. Kamath N, Tsvankin I. Elastic full-waveform inversion for VTI media: Methodology and sensitivity analysis. *Geophysics*. 2016;81(2):C53-C68.

    doi: 10.1190/geo2014-0586.1

27. Li V, Tsvankin I, Alkhalifah T. Analysis of RTM extended images for VTI media. *Geophysics*. 2016;81(3):S139-S150.

    doi: 10.1190/geo2015-0384.1

28. Luo C, Ba J, Carcione JM. A hierarchical prestack seismic inversion scheme for VTI media based on the exact reflection coefficient. *IEEE Trans Geosci Remote Sens*. 2022;60:1-16.

    doi: 10.1109/TGRS.2021.3140133

29. Song C, Alkhalifah T, Waheed UB. Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks. *Geophys J Int*. 2021;225(2):846-859.

    doi: 10.1093/gji/ggab010

30. Ristow D. Migration in transversely isotropic media using implicit finite-difference operators. *J Seism Explor*.

1999;8:39-56.

31. Han Q, Wu RS. A one-way dual-domain propagator for scalar qP-waves in VTI media. *Geophysics*. 2005;70(2):D9-D17.

doi: 10.1190/1.1884826

32. Bale RA. Phase-Shift Migration and the Anisotropic Acoustic Wave Equation. In: *Proceedings of the 69th EAGE Conference and Exhibition Incorporating SPE EUROPEC 2007*; June 11-14, 2007; London, UK. European Association of Geoscientists & Engineers; 2007:cp-27-00104.

doi: 10.3997/2214-4609.201401522

33. Fei TW, Liner CL. Hybrid Fourier finite-difference 3D depth migration for anisotropic media. *Geophysics*. 2008;73(2):S27-S34.

doi: 10.1190/1.2828704

34. Bakker PM. A stable one-way wave propagator for VTI media. *Geophysics*. 2009;74(5):WB3-WB17.

doi: 10.1190/1.3196818

35. Amazonas D, Aleixo R, Schleicher J, *et al*. Anisotropic complex Padé hybrid finite-difference depth migration. *Geophysics*. 2010;75(2):S51-S59.

doi: 10.1190/1.3337317

36. Salcedo M, Novais A, Schleicher J, *et al*. Optimization of the parameters in complex Padé Fourier finite-difference migration. *Geophysics*. 2017;82(3):S259-S269.

doi: 10.1190/geo2016-0324.1

37. Liu LN, Zhang JF. Wave equation prestack depth migration method in 3D VTI media. *Chin J Geophys*. 2011;54(6):844-855.

doi: 10.1002/cjg2.1667

38. Vyas M, Wang X, Etgen J. One-way TTI propagation: Is it still relevant? In: *SEG Technical Program Expanded Abstracts 2015*; Society of Exploration Geophysicists; 2015:3986-3990.

doi: 10.1190/segam2015-5915894.1

39. Alshuhail AA, Verschuur DJ. Robust estimation of vertical symmetry axis models via joint migration inversion: Including multiples in anisotropic parameter estimation. *Geophysics*. 2019;84(1):C57-C74.

doi: 10.1190/geo2017-0856.1

40. Wu B, Wu RS, Gao J. Preliminary Investigation of Wavefield Depth Extrapolation by Two-Way Wave Equations. *Int J Geophys*. 2012;2012:968090.

doi: 10.1155/2012/968090

41. You J, Wu RS, Liu X, *et al*. Two-way wave equation-based depth migration using one-way propagators on a bilayer sensor seismic acquisition system. *Geophysics*. 2018;83(3):S271-S278.

doi: 10.1190/geo2017-0172.1

42. You J, Pan N, Liu W, *et al*. Efficient wavefield separation by reformulation of two-way wave-equation depth-extrapolation scheme. *Geophysics*. 2022;87(4):S209-S222.

doi: 10.1190/geo2021-0629.1

43. You JC, Zhang GL, Huang XG, *et al*. Internal-multiple-elimination with application to migration using two-way wave equation depth-extrapolation scheme. *Pet Sci*. 2025;22(1):178-192.

doi: 10.1016/j.petsci.2024.06.021

44. Li A, Zhu X, Liu T, *et al*. One-step sparse-matrix method for full wave equation depth migration. *Geophysics*. 2024;89(6):S405-S413.

doi: 10.1190/geo2023-0372.1

45. Sandberg K, Beylkin G. Full-wave-equation depth extrapolation for migration. *Geophysics*. 2009;74(6):WCA121-WCA128.

doi: 10.1190/1.3202535

46. You J, Cao J. Full-wave-equation depth extrapolation for migration using matrix multiplication. *Geophysics*. 2020;85(6):S395-S403.

doi: 10.1190/geo2019-0323.1

47. You J, Sun R, Huang X, Huang J. DeepImaging: Deep neural networks driven full waveform migration. *IEEE Trans Geosci Remote Sens*. 2026;64:5900719.

doi: 10.1109/TGRS.2025.3648341

48. Sun R, You J, Li Z, Hu H. Viscoacoustic Full Wavefield Migration and Its Application. *IEEE Trans Geosci Remote Sens*. 2025;63:5914517.

doi: 10.1109/TGRS.2025.3574991

49. Alkhalifah T. Traveltime computation with the linearized eikonal equation for anisotropic Media. *Geophys Prospect*. 2002;50(4):373–382.

doi: 10.1046/j.1365-2478.2002.00322.x

50. Liu YZ, Wang GY, Dong LG, *et al*. Joint inversion of VTI parameters using nonlinear traveltime tomography. *Chin J Geophys*. 2014;57(10):3402-3410. [In Chinese].

doi: 10.6038/cjg20141026

## Appendix A

### Appendix A.1. A frequency–wavenumber (FK) filter for removing pseudo-S-waves

To suppress the evanescent waves and pseudo-S-waves in the FK, the vertical wavenumber is reformulated as follows:

$$\begin{cases} k_z^2 = \dfrac{\omega^2}{v_p^2}\left(\dfrac{A}{B}\right) \\ A = \omega^2 - v_n^2(1+2\eta)k_x^2 \\ B = \omega^2 - 2v_n^2\eta k_x^2 \end{cases} \tag{A1}$$

When both parameters A > 0 and B > 0, the resulting wavefield corresponds to a propagating P-wave. In contrast, if A < 0 and B < 0, the solution describes a pseudo-S-wave. To suppress these undesired pseudo-S-wave components, the associated one-way propagation operator can be effectively muted by assigning it a value of zero under the condition A < 0 and B < 0.

To evaluate the effectiveness of removing pseudo-S-waves, a velocity model $v_p(z,x) = 1500 + 2x + 2z$ was generated, as shown in **Figure A1**, with anisotropic parameters $\varepsilon = 0.6$ and $\delta = 0$. The modeled wavefields with and without removing pseudo-S-waves are shown in **Figure A2**. As indicated with red arrows in **Figure A1**, the pseudo-S-waves can be observed without using the FK filter, while clear P-waves are shown without the pseudo-S-waves when using the FK filter.

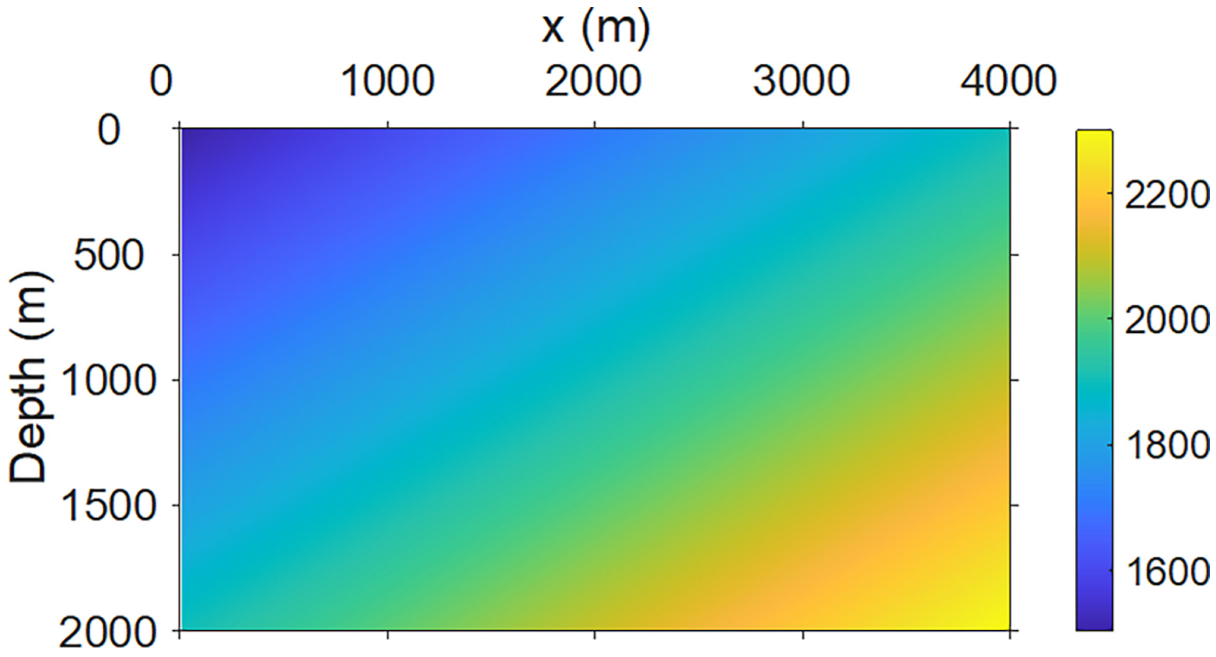



**Figure A1.** Velocity model calculated using $v_p(z,x) = 1500 + 2x + 2z$

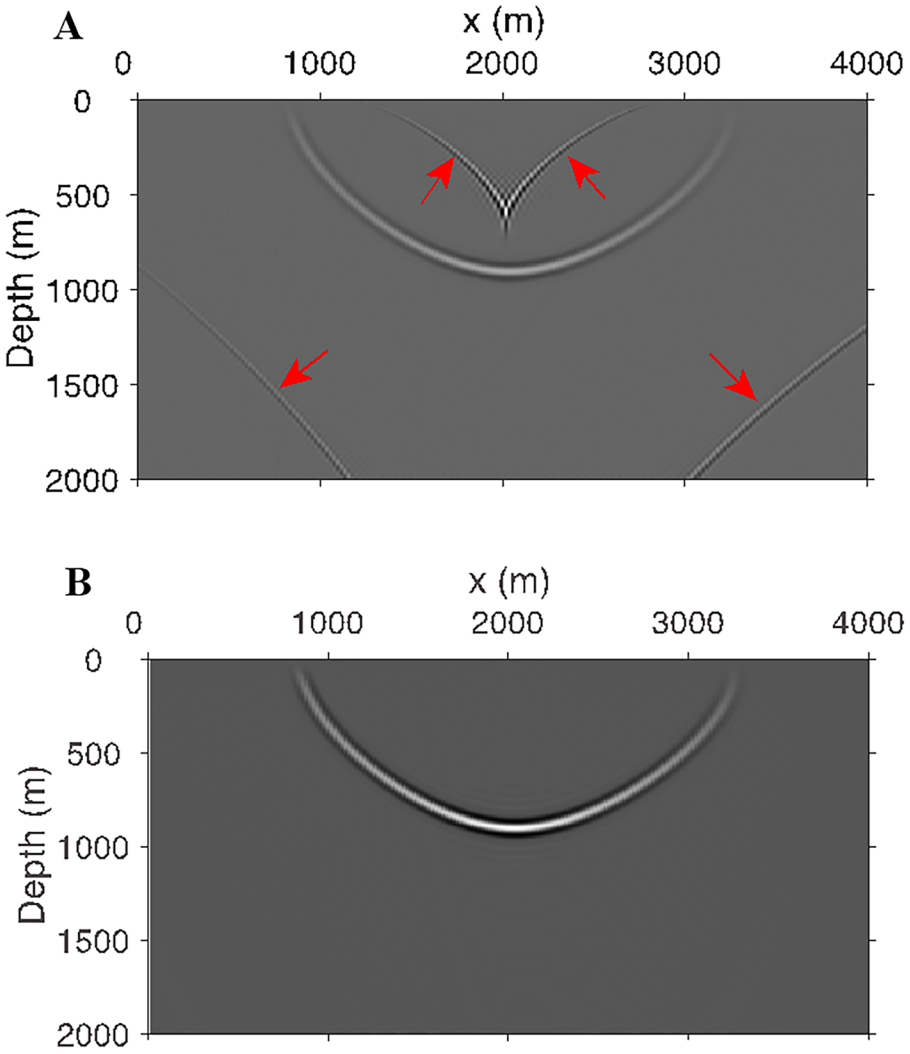



**Figure A2.** Modeled wavefield (A) without removing pseudo-S-waves and (B) with pseudo-S-waves removed. Red arrows indicate pseudo-S-waves.

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Suppression of surface waves in seismic data using an adaptive time–frequency–wavenumber filter

**Xuejie Gao[1,2], Yong Wang[1,2]*, Youjuan He[3], Zhili Chen[1,2], Zerun Nian[4], and Lang Yang[4]**

[1]Department of Geophysical Exploration, School of Geophysics and Petroleum Resources, Yangtze University, Wuhan, Hubei, China

[2]Key Laboratory of Oil and Gas Resources and Exploration Technology, Yangtze University, Wuhan, Hubei, China

[3]Research Institute of Exploration and Development, Northwest Oilfield Company, SINOPEC, Urumqi, Xinjiang Uygur Autonomous Region, China

[4]Exploration and Development Research Institute, Geophysical Research Department, PetroChina Tarim Oilfield Branch, Korla, Xinjiang Uygur Autonomous Region, China

## Abstract

Surface waves are a prevalent form of coherent interference in seismic recordings, characterized by low frequency, high energy, and slow velocity, which can significantly affect seismic data processing. Currently, surface wave suppression is primarily achieved by leveraging the differences between surface waves and signal waves in the frequency, amplitude, and wavenumber domains. Among these methods, frequency–wavenumber (FK) filtering is widely used. However, it requires manual selection of regions during filtering, which becomes challenging when processing large volumes of seismic data. Therefore, adaptive surface-wave suppression methods are essential for practical data processing. FK filtering transforms data into the FK domain to suppress linear noise. However, because surface waves are dispersive and exhibit partial nonlinearity, FK filtering alone is insufficient to fully suppress them. To address these issues, this study introduces a temporal dimension into the FK domain, leading to the development of the time–frequency–wavenumber (TFK) transformation. This transformation further separates surface and signal waves in the time domain, thereby mitigating dispersion. Moreover, it facilitates adaptive filtering of seismic data across different time intervals in the FK domain. By cross-correlating FK data from different time periods, adaptive filters were derived for each interval, which were then applied to obtain filtered seismic records. Comparisons between synthetic and real-world data demonstrated that our approach effectively suppresses surface waves while preserving the relative amplitude characteristics of signal waves.

***Keywords:*** Seismic noise suppression; Surface-wave suppression; Frequency dispersion; Frequency–wavenumber filter; Adaptive time–frequency–wavenumber filtering; Cross-correlation

***Corresponding author:**
Yong Wang
(cdwangyong@yangtzeu.edu.cn)

**Citation:** Gao X, Wang Y, He Y, Chen Z, Nian Z, Yang L. Suppression of surface waves in seismic data using an adaptive time–frequency–wavenumber filter. *J Seismic Explor.* 2026;35(2):025450103. doi: 10.36922/JSE025450103

**Received:** November 5, 2025

**Revised:** February 9, 2026

**Accepted:** February 12, 2026

**Published online:** April 27, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Noise suppression holds significant importance in petroleum seismic exploration, as it exerts a notable influence on subsequent processing. Seismic data are predominantly affected by diverse wave types, with surface waves being prominent. Surface waves feature a fan-shaped distribution, characterized by high energy, low frequency, slow velocity, and frequency dispersion.[1,2] Their presence conceals seismic reflections that are vital for mapping subsurface structures. This concealment effect not only complicates the extraction of valuable geological information but also heightens the risk of misinterpretation, thereby directly affecting the accuracy of hydrocarbon reservoir evaluations and geological hazard assessments. Conventional methods aimed at alleviating this effect often encounter challenges. They struggle to effectively separate and suppress surface waves while safeguarding the integrity of seismic signals.[3]

Surface waves can be grouped into Rayleigh and Love waves. A Rayleigh wave is a seismic surface wave produced by the interference between surface longitudinal and transverse waves. On the other hand, Love waves originate when S-waves travel along layered interfaces. Both kinds of waves propagate along the surface.[4] In the realm of petroleum seismic exploration[5,6], these waves are usually suppressed. However, in engineering seismic exploration, they are used for tasks such as determining crack morphology and detecting in-situ damage.[7–9] In petroleum seismic exploration, vertical-component geophones are commonly used, which results in the reception of Rayleigh waves. Therefore, this research centers on Rayleigh surface waves.

Traditional approaches for attenuating surface waves primarily include high-pass filters, frequency–wavenumber (FK) filters, and time-intercept slowness filters. Although these approaches are somewhat efficient and effective, each has distinct drawbacks.[10] As a result, in recent years, a variety of surface wave analysis methods have emerged, such as the K–L transform, the slant FK filter, and deep learning techniques.[11–13] For example, Jin *et al.*[14] put forward a new inverse dispersion method. This method addresses the nonlinearity caused by surface wave dispersion, enabling subsequent surface wave suppression through FK filtering. Liu *et al.*[15] proposed a method integrating a blind-trace scheme and a self-supervised neural network. Its purpose is to eliminate aliased cone-shaped surface waves in seismic data. Lv *et al.*[3] introduced a nonsubsampled Shearlet transform method. First, it conducts a preliminary separation of surface waves and effective waves. Then, the shearlet transform is used to decompose the data into multiple sub-bands. After that, correlation calculations are performed to adaptively select the effective sub-bands. Lastly, an inverse transform is applied to the selected sub-bands. In response to the instability of the U-Net and related issues, Son *et al.*[16] proposed the blind horizon network and dual-model self-supervised selective learning. Yuanyuan *et al.*[10] presented a novel model-based, data-driven method for the attenuation of surface waves. This approach uses surface wave dispersion curve analysis and a joint inversion that combines the genetic algorithm and the conjugate gradient algorithm to develop a precise surface wave model. Subsequently, the model data are subtracted from the seismic data to enhance the signal-to-noise ratio.

Recent research indicates that the focus has mainly been on adaptivity and data analysis. Drawing on traditional methods, this study explores adaptive strategies to further advance the field's development. Yablokov *et al.*[17] developed a novel approach for determining static corrections for the hodograph rectification of surface waves in the time–frequency (TF) domain. This method also involves preprocessing, in which the reflected wave signals are removed in advance.

The FK filtering technique differentiates between surface waves and reflected waves by exploiting their frequency and apparent velocity differences. In specific situations, it can efficiently reduce linear noise.[18,19] Nevertheless, when surface and reflected waves show aliasing in the FK domain, or when the strong dispersion of surface waves diminishes their linear features, the performance of FK filtering is notably reduced. Numerous researchers have introduced the time–frequency–wavenumber (TFK) method by enhancing certain attributes to improve upon the original FK filtering approach. For instance, Askari and Siahkoohi[20] proposed the TFK method, which incorporates time information into the FK framework. By establishing different suppression areas in the FK domain at various time intervals, this method boosts surface wave suppression. Liu and Fomel[21] proposed the l-TFK method. This method first applies the TFK transformation to the seismic data and then decomposes it using regularization techniques. Nonetheless, TFK filtering requires manual definition of a rejection zone at each time point. This poses a challenge when dealing with a large amount of seismic data. To address this issue, Wang and Wang[22] proposed a polarization filtering approach based on the TFK method. This method creates an analytic signal by combining the horizontal and vertical amplitudes and automatically determines the rejection zone using the inverse of the ellipticity, thereby suppressing surface waves. However, this method is not suitable for data that comprise only vertical components. Wu *et al.*[23] proposed an automatic

data selection method for surface wave imaging, which is based on the energy distribution characteristics in the FK domain. Sun *et al.*[24] employed FK filtering to suppress single-frequency noise and linear noise and evaluated its excellent amplitude-preserving ability.

This paper presents an adaptive TFK filter to address the challenges of automatic filter configuration and the overlap between surface and effective waves, thereby improving surface wave suppression. In traditional FK filtering, accurately picking the surface wave region is a crucial step. However, conventional picking methods often fall short when dealing with the overlapping areas of surface and effective waves. These methods are unable to precisely delineate the boundaries of surface waves in such complex regions, resulting in incomplete suppression and potential distortion of the effective wave components. To address this limitation, an adaptive TFK filter was proposed. Initially, the TFK method was applied to seismic data containing surface waves, thereby acquiring data in the TFK domain. After that, cross-correlation operations were performed on the data in the FK domain across different time dimensions to obtain an adaptive filter. This adaptive filter could intelligently identify and separate surface wave components, even in overlapping areas, providing a more accurate filtering effect than traditional methods. Finally, the filters were applied to the FK domain data at various time points. Subsequently, an inverse TFK transformation was performed, resulting in the data with surface waves suppressed. Our approach can be used for both synthetic data and actual seismic records. By comparing it with the traditional FK filtering method, the remarkable advantages of the adaptive TFK filter were demonstrated.

## 2. Materials and methods

### 2.1. Time–frequency–wavenumber transformation

The TFK transformation is a signal-processing approach for extracting TFK information from seismic data. This methodology extends the FK transformation by integrating time information. Motivated by the S-transform-based ground-roll suppression approach validated by Askari and Siahkoohi[20], we applied the S-transform to each seismic trace within a shot gather to obtain a frequency-adaptive, multi-resolution TF representation. This frequency-dependent resolution is particularly suitable for land seismic data, where surface waves are typically non-stationary and exhibit strong dispersion—often characterized by low frequencies and long durations. In contrast, reflection events tend to manifest as more transient, higher-frequency features. Consequently, fixed-window TF representations, such as the short-time Fourier transform, may suffer from inherent trade-offs between temporal and spectral resolution. Furthermore, the S-transform preserves a physically meaningful Fourier frequency axis and retains the full complex spectrum, enabling seamless integration with subsequent spatial Fourier transforms. This compatibility naturally supports the correlation-based operations performed on constant-frequency (FK) slices in our TFK-domain workflow.

The S-transform[25] is a method that represents a signal as a 2D function of frequency and time, enabling simultaneous observation of its characteristics in both domains. It employs a window function in the frequency and time domains to locally weight the signal and transform it at different times and frequencies. This results in a TF representation that shows the signal's energy distribution across time and frequency. The equation is as follows:

$$S(\tau,f)=\frac{|f|}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} r(t)e^{\left[-\frac{1}{2}(\tau-t)^2 f^2\right]}e^{-i2\pi ft}dt \tag{1}$$

where $r(t)$ is the seismic data of a single trace, $t$ and $\tau$ represent the time variables for the single-trace seismic data and the TF spectral data, respectively, $f$ represents the frequency, $i$ represents the imaginary part of the signal, and $S(\tau,f)$ represents the TF spectral data of the seismic signal.

After performing the S-transform on each trace of the seismic data, $S(\tau,f,c)$ is obtained, where $c$ is the spatial variable of the seismic data, representing the offset distance. A Fourier transform is performed in the spatial direction to convert spatial information into wavenumber information. The equation is as follows:

$$\overline{S}(\tau,f,k)=\frac{1}{2\pi}\int S(\tau,f,c)e^{-i2\pi kc}dc \tag{2}$$

where $\overline{S}(\tau,f,k)$ represents the TFK transformation, which includes the results of time, frequency, and wavenumber, and $k$ denotes the spatial wavenumber.

Since spatial sampling is often irregular in practical seismic data acquisition, appropriate adjustments to **Equation 2** are required when performing the spatial Fourier transform. In seismic data processing, irregular sampling can lead to spectral leakage and aliasing artifacts in the FK domain.[26] To mitigate these effects, we adopted an antileakage Fourier transform-based weighting scheme by introducing an integral weight $w(c_j)$ on irregular samples $\{c_j\}_{j=1}^{N}$, and discretized **Equation 2** as:

$$\begin{aligned}&\overline{S}(\tau,f,k)=\frac{1}{\Delta X}\sum_{j=1}^{N}\left[w(c_j)\overline{S}(\tau,f,c_j)e^{-2\pi ikc_j}\right],\\&\overline{S}(\tau,f,c_j)=\overline{S}(\tau,f,k)e^{2\pi ikc_j}\end{aligned} \tag{3}$$

where $w(c_j)$ is the integral weight, which is the reciprocal

of the sampling density; ΔX is a normalization factor involving the summation of all the sampled integral weights; and $N$ is the number of input points.

### 2.2. Adaptive time–frequency–wavenumber filter

The adaptive TFK filtering technique suppresses surface waves in the TFK domain. After dispersion inversion processing, surface waves exhibit linear characteristics, and their energy is typically significantly stronger than that of effective waves. According to the convolution theorem, cross-correlation in the time domain corresponds to conjugate multiplication in the frequency domain. To accurately identify the distribution range of surface waves in the FK domain, this study computed FK characteristics corresponding to different time points to construct a discriminative weighting function. Specifically, by performing multiplicative operations on signals from adjacent time points in the TF domain, the cross-correlation in the time–space domain was calculated, thereby generating the weighting coefficients required for subsequent adaptive filtering. The specific equation is as follows:

$$w(\tau,f,k)=\sum_{i=\tau-\delta}^{\tau+\delta}\left|\overline{S}(\tau,f,k)\cdot\mathrm{conj}\left(\overline{S}(i,f,k)\right)\right| \quad (4)$$

where $w(\tau,f,k)$ is a correlation-based attribute used to characterize the relative strength and likelihood of surface wave energy at the TFK-domain coordinate $(\tau,f,k)$; $\tau$ denotes a discrete time-sample index along the time dimension of the TFK transform, serving as the center of a local correlation window; and $\delta$ denotes the half-window length in samples (an integer index offset after rounding), which controls the number of neighboring time samples involved in **Equation 4**. Accordingly, the full window length corresponds to $(2\delta+1)$ samples, i.e., $(2\delta+1)\Delta t$ in time with $\Delta t=1/f_s$. In this study, the window length was designed to cover approximately four cycles of the dominant surface wave frequency $f_0$, leading to **Equation 5**.

$$\delta=\mathrm{round}\left(\frac{2f_s}{f_0}\right) \quad (5)$$

where $f_s$ is the sampling frequency, $f_0$ is the dominant frequency of the surface wave, and $\mathrm{round}(\cdot)$ denotes rounding to the nearest integer to ensure $\delta$ is an integer sample offset.

After performing the correlation calculations, $w(\tau,f,k)$ still contains regions of effective waves, necessitating additional refinement of $w(\tau,f,k)$. Since the inverse-dispersed surface waves exhibit linear characteristics and have the highest energy, a linear scan of $w(\tau,f,k)$ was conducted to identify the region with peak energy as the surface wave area, while other regions were classified as non-surface wave areas. The equation is as follows:

$$w_0(k,f)=\begin{cases}1 & |af-k|<\dfrac{4\pi f_o}{v}\\ 0 & other\end{cases} \quad (6)$$

where $w_0$ is the computational region, $a$ is the surface wave slope obtained from energy scanning, and $v$ is the velocity of surface waves. It should be noted that the surface wave slope *a* here is not a globally physically precise surface wave propagation velocity, but a locally valid approximation within short-time windows in the TFK domain. It is used only to initially constrain the candidate surface wave region, and local slope fluctuations have a limited impact on the final separation result.

A computational region is obtained through linear scanning and is then multiplied by the stacking area for further refinement. The refined stacking area is subsequently subjected to inverse computation to derive the final filtering region. Finally, filtering is performed in the TFK domain.

$$s_0(t,c)=itfk\left(\sum_{j=1}^{T}\left(\overline{S}(j,f,k)\cdot(1-\omega)\right)\right) \quad (7)$$

where $itfk$ is the inverse TFK, and $s_0(t,c)$ represents the data after surface wave suppression.

To clarify the overall processing procedure, the main steps of the proposed adaptive TFK filter are summarized in Figure 1.

As shown in Figure 1, the input shot gather in the time-offset domain is first transformed into the TFK spectrum by combining the S-transform in the TF domain with the spatial Fourier transform along the offset direction. Then, for each frequency and wavenumber, a time-windowed cross-correlation is computed between adjacent time slices of the TFK spectrum. This operation enhances the coherent and energetic components, primarily surface waves, and yields a correlation attribute that characterizes the strength of surface wave energy in the TFK domain. The correlation attribute is subsequently normalized to construct an adaptive TFK filter, which delineates the surface wave cone while preserving the reflection-dominated region. Finally, the adaptive filter is applied to the TFK spectrum, and the denoised shot gather is obtained in the time-offset domain by performing the inverse spatial Fourier transform together with the inverse S-transform.

## 3. Results

In this study, both synthetic and field seismic datasets were used to systematically compare the conventional FK filtering technique with the proposed adaptive TFK filtering

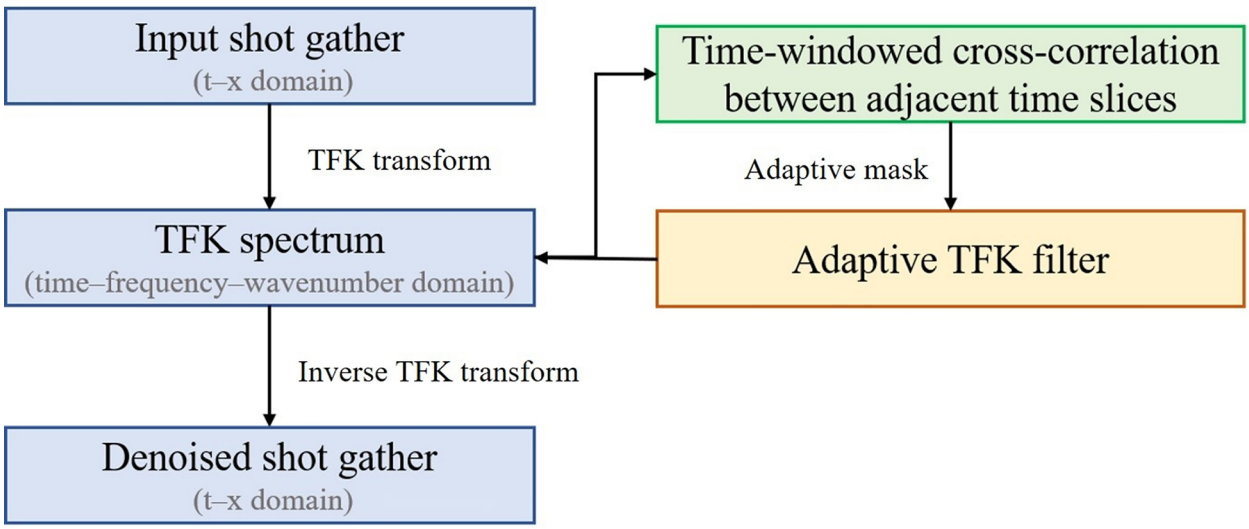



**Figure 1.** Flowchart of the proposed time–frequency–wavenumber (TFK) adaptive filter for surface-wave suppression

approach. The primary objective of this comparative analysis was to evaluate the relative performance of the two methods in suppressing dispersive surface wave energy while preserving useful reflection events. By applying both algorithms to the same dataset and examining their results in both the time and FK domains, the study aims to assess their differences in identifying surface wave components, maintaining the fidelity of practical reflections, and improving the overall signal-to-noise ratio. The insights gained from this comparison are intended to provide valuable guidance for the further optimization of surface-wave suppression strategies in seismic data processing.

### 3.1. Comparison of simulation data

In the synthetic experiment, a single-shot seismic record was constructed by superimposing a dispersive surface wave onto three hyperbolic reflection events. The reflected signals were assigned a dominant frequency of 20 Hz, while the surface wave was designed with a lower dominant frequency of approximately 10 Hz, mimicking a typical near-surface seismic scenario where low-frequency surface energy interferes with higher-frequency reflections. The sampling frequency was set to 200 Hz, and the total record length was 3.5 s, comprising 50 traces. The synthetic dataset was subsequently transformed into the FK domain using the FK transform (Figure 2).

As illustrated in Figure 2A, the time-domain single-shot record showed a nearly linear surface wave event that overlapped with the reflection signals, resulting in noticeable dispersion. The corresponding FK spectrum, shown in Figure 2B, revealed that the surface wave energy was distributed along a straight trajectory in the FK plane and partially overlapped with the effective reflection energy band, exhibiting the typical low-frequency, high-amplitude characteristics of ground roll. This clear visualization allows for a direct assessment of the interaction between the surface and reflected waves in the spectral domain. Meanwhile, due to spatial sampling limitations in the dataset, pronounced spatial aliasing occurred, further exacerbating the overlap between the surface wave and effective wave energy in the FK spectrum. Consequently, the direct application of conventional FK filtering is prone to inadequate wave separation, leading to potential processing inaccuracies.

Building on this, a TFK transform was applied to the data to obtain a TFK spectrum, enabling a more detailed investigation of how surface wave energy evolves. FK spectra corresponding to different time instants were extracted to capture the temporal variability of the surface wave distribution. Because the FK spectrum is symmetric about the frequency axis, only the positive-frequency portion is presented here for clarity and conciseness.

Figure 3 illustrates the FK spectra obtained at three representative time instants: 0.5 s, 2 s, and 3 s. As shown, the spectra at 0.5 s and 3 s were dominated almost entirely by surface wave energy, while the spectrum at 2 s revealed a coexistence of both surface and reflection waves. This observation highlights the temporal overlap and complex interaction between dispersive ground roll and effective reflections, which pose a significant challenge for accurate separation and suppression in seismic data processing.

In this research, the energy of the signal waves was intentionally amplified relative to that of the surface waves, primarily to verify the stability of the proposed method

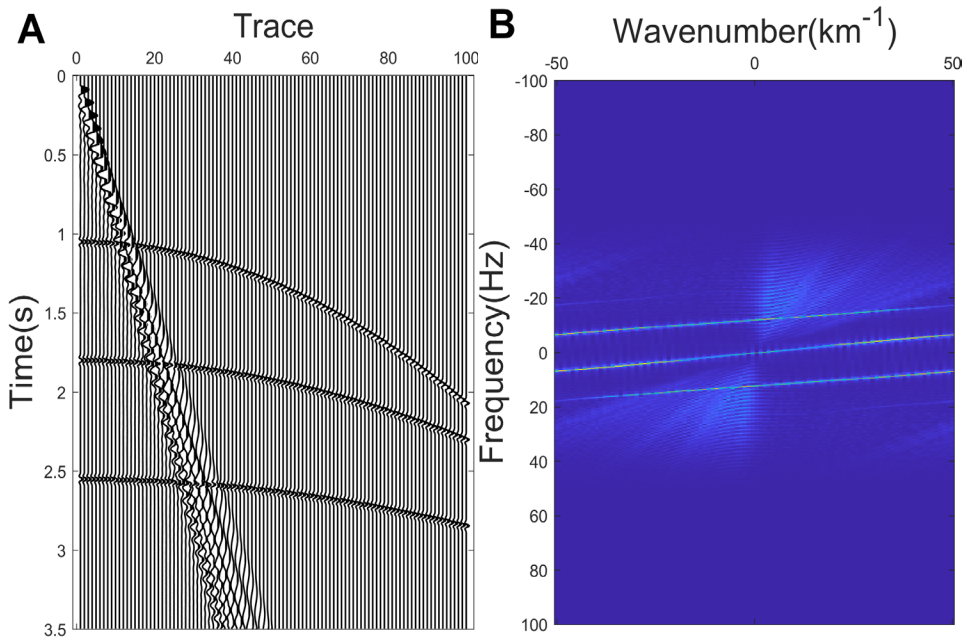



**Figure 2.** Generated data and corresponding frequency–wavenumber (FK) spectrum: (A) single-shot data and (B) FK spectrum of the synthetic data

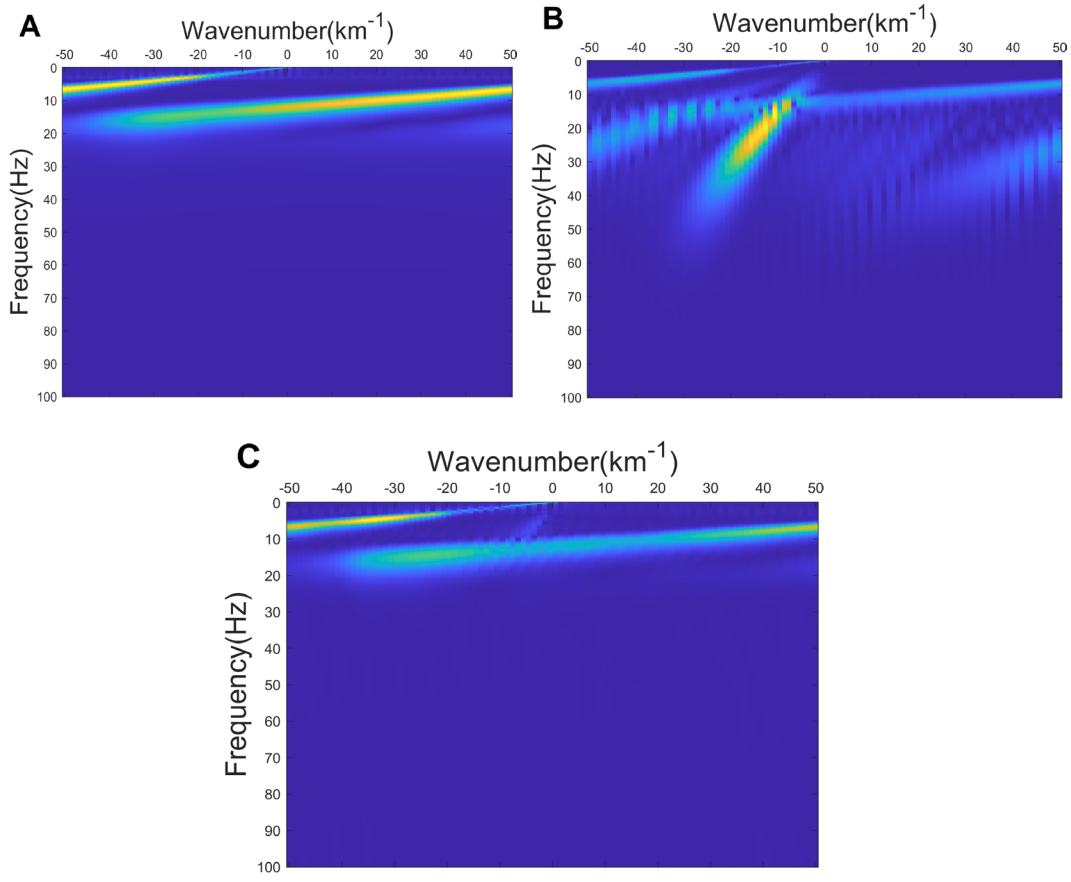



**Figure 3.** Frequency–wavenumber (FK) spectra at different time instants in the time–frequency–wavenumber transform. The FK spectral data at (A) 0.5 s, (B) 2 s, and (C) 3 s.

and to demonstrate that it relies not on energy magnitude but on energy distribution characteristics. As illustrated in Figure 3B, the energy of the signal waves was significantly higher than that of the surface waves, to the extent that it distorted the surface waves to a certain degree—this objectively increased the difficulty of extracting the surface wave region. Because surface waves are inherently dispersive, their FK spectra vary with time as the phase velocity changes with frequency. Although the distribution of surface wave energy exhibited slight differences across distinct time instants, the variations remained relatively minor within neighboring time windows. This property enables the implementation of a cross-correlation-based detection algorithm to identify regions dominated by surface wave energy, thereby enabling localized and adaptive suppression within the FK domain. In this research, we constructed a fan filter based on the surface wave velocity for FK filtering and then compared the results with those obtained via adaptive TFK filtering (Figure 4).

Figure 4 compares the performance of the conventional FK filter and the proposed adaptive TFK filter for both noise-free (Figure 4A–C) and noisy synthetic data (Figure 4D–F). In noise-free cases, the FK filter partially reduced surface wave energy while preserving the general trend of the reflection events. However, in regions where surface waves and reflections overlapped, the FK filter failed to achieve precise separation, leading to blurred reflection boundaries and residual low-frequency energy. This limitation is mainly due to the fixed-window nature of the FK domain, which limits its joint TF and spatial resolution and makes it difficult to completely suppress dispersive tail waves.

In noisy synthetic data cases, FK filtering showed strong residual surface wave energy and heavily contaminated reflections. In contrast, the adaptive TFK filter largely suppressed the coherent surface waves, and the remaining energy was mainly weak, incoherent background noise. The adaptive TFK filtering method introduced a joint TF analysis mechanism that dynamically adjusted the filtering response according to the local spectral characteristics of the data. This adaptive framework enables the algorithm to achieve a superior balance between waveform preservation and noise suppression, effectively reducing surface wave energy while maintaining the integrity of the reflection signals. Such improvements highlight the advantages of incorporating temporal adaptivity into surface-wave suppression procedures. Additionally, the residual random noise can be further suppressed by standard random-noise suppression techniques.

To further evaluate the surface-wave suppression performance of the two methods, FK spectra at different time instants after filtering were analyzed and compared, as illustrated in Figure 5. Figure 5A–C shows the results obtained from the traditional FK filter, while Figure 5D–F displays the spectra derived from the adaptive TFK filtering approach. In the original dataset, the FK spectra at 0.5 s and 3 s were dominated by surface wave energy, whereas the spectrum at 2 s exhibited a mixture of surface and reflected waves.

As shown in Figure 5B, the conventional FK filtering method failed to eliminate the surface wave components. The suppression effect was primarily confined to the 0–10 Hz frequency range, beyond which substantial residual energy persisted. In regions where surface and reflection signals overlapped, the FK filter was largely ineffective, resulting in incomplete suppression.

By contrast, the adaptive TFK filtering method demonstrated superior performance: it effectively distinguished overlapping components and provided more thorough suppression of surface wave energy, even within mixed zones. Notably, in the FK spectrum at 2 s, the TFK approach retained the reflection signal to a greater extent while simultaneously suppressing the interfering surface waves. This outcome highlights that the proposed method exhibits excellent stability in surface wave region extraction and superior performance in surface wave suppression.

Figure 6 compares a single seismic trace (the 50th channel) before and after the application of the two filtering methods. As depicted, the FK method partially suppressed surface wave energy, but a substantial residual persisted at low frequencies. In contrast, the TFK approach demonstrated stronger suppression, with the surface wave component nearly eliminated from the record.

Nevertheless, a slight reduction in the amplitude of the reflection events was observed after TFK processing, indicating some degree of signal suppression. This issue can be mitigated by applying relative normalization or amplitude-compensation techniques during post-processing. Overall, the adaptive TFK filtering method achieved a more balanced trade-off between surface wave suppression and signal preservation, resulting in seismic records with improved clarity and enhanced signal-to-noise ratio. These findings confirm that incorporating TF adaptivity into the FK framework provides a more robust and flexible solution for addressing surface wave interference in seismic data processing.

In summary, the experimental results clearly demonstrate that the proposed adaptive TFK filtering method outperforms the traditional FK filtering technique in surface wave suppression. It effectively distinguishes and suppresses surface wave energy even within regions

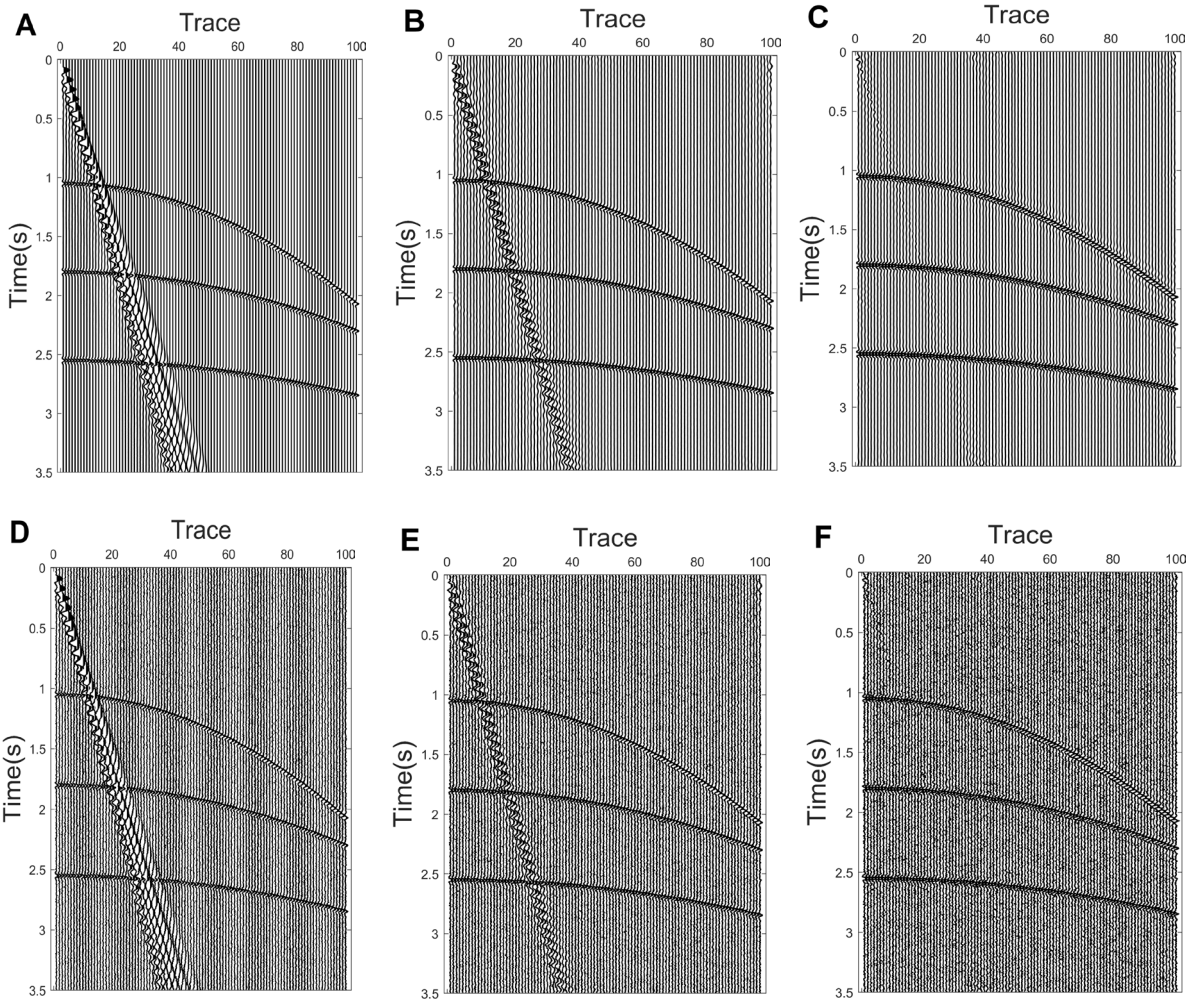



**Figure 4.** Comparison of frequency–wavenumber (FK) surface wave suppression and time–frequency–wavenumber (TFK) adaptive surface wave suppression: (A) Original noise-free synthetic data; (B) FK filtering results of (A); (C) adaptive TFK filtering results of (B); (D) original data with random noise; (E) FK filtering results of (D); and (F) adaptive TFK filtering results of (D)

affected by spatial aliasing and spatial overlap, while maintaining the integrity of the reflection signals. Beyond its improved suppression capability, the TFK approach enhances both the signal-to-noise ratio and the temporal–spatial resolution of the seismic data. These advantages underline its potential as a practical and powerful tool for seismic data processing under complex near-surface conditions, offering new perspectives for the development of adaptive, high-resolution filtering algorithms in applied geophysics.

### 3.2. Comparison of field data

To further evaluate the advantages of the proposed approach, two representative field seismic datasets were selected to compare the performance of the conventional FK filtering technique with the proposed adaptive TFK filtering method in suppressing surface waves. Both datasets have a record length of 2 s and a sampling frequency of 500 Hz. Figure 7 summarizes the raw seismic records from the two datasets, along with their corresponding one-sided FK spectra, which served as the basis for the subsequent comparison of filtering performance.

As illustrated in Figure 7, surface wave energy dominated both field seismic datasets, with its intensity being significantly higher than that of effective reflection wave energy. Notably, the spatio-temporal distributions of surface waves differed between the two datasets: in Dataset 1 (Figure 7A), strong surface wave energy persisted throughout the entire time window, whereas in Dataset 2 (Figure 7C), it was concentrated locally and gradually attenuated over time. Further analysis of the unilateral FK spectrum revealed a distinct overlap

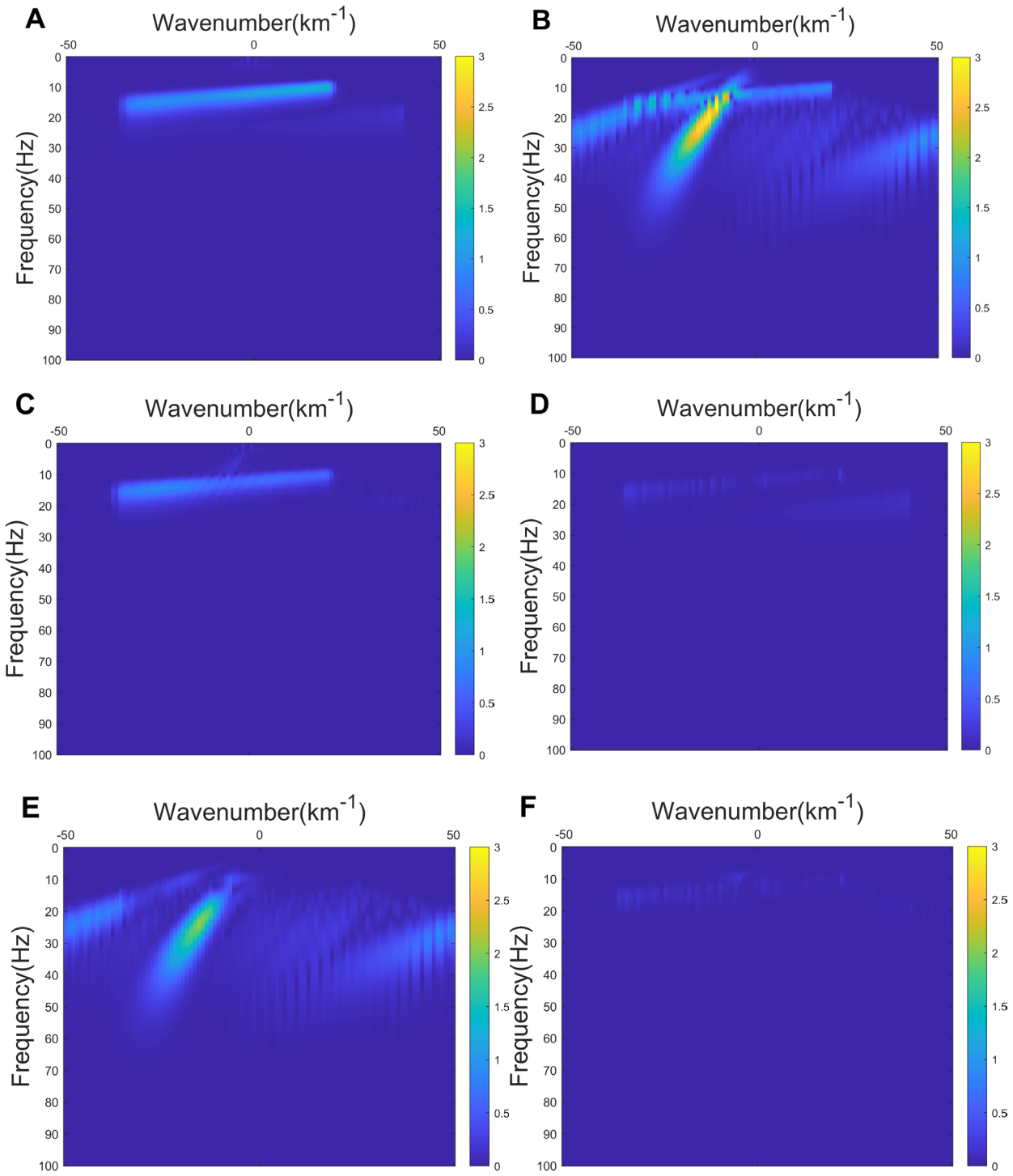



**Figure 5.** Distinct frequency–wavenumber (FK) spectra after suppression using two methods: (A–C) the FK spectrum at (A) 0.5 s, (B) 2 s, and (C) 3 s after FK filtering; and (D–F) the FK spectrum at (D) 0.5 s, (E) 2 s, and (F) 3 s after adaptive time–frequency–wavenumber filtering

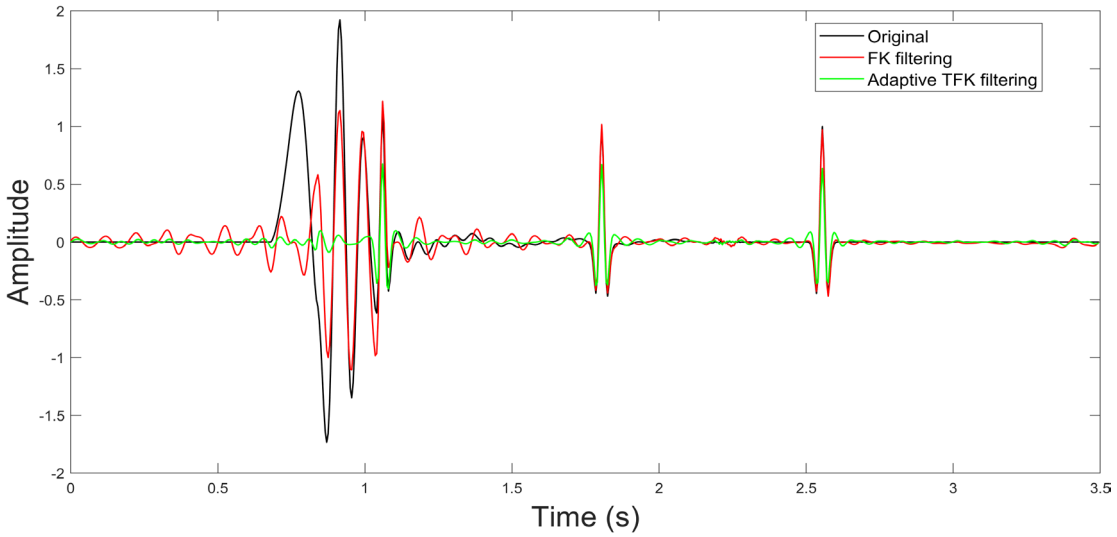



**Figure 6.** Comparison of frequency–wavenumber (FK) spectra after suppression: original data (black line), FK filtering (red line), and adaptive time–frequency–wavenumber (TFK) filtering (green line)

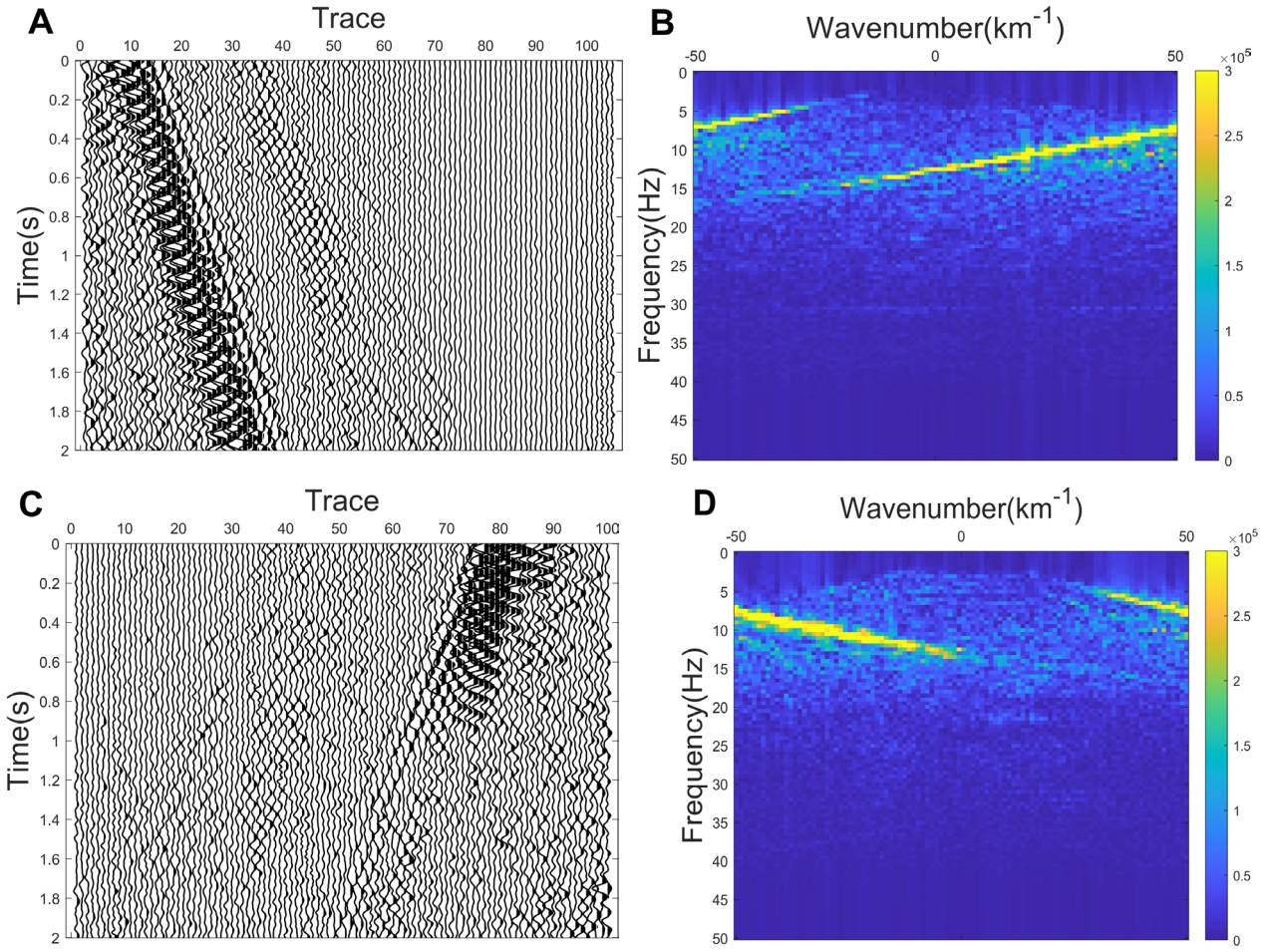



**Figure 7.** Two field seismic datasets and their corresponding one-sided frequency–wavenumber (FK) spectra. (A) Field seismic data (Dataset 1). (B) One-sided FK spectrum of Dataset 1. (C) Field seismic data (Dataset 2). (D) One-sided FK spectrum of Dataset 2.

between the dominant frequency bands of surface waves and reflection waves. Under such circumstances, it is challenging to delineate a clear separation cone in the FK domain, which renders traditional FK filtering ineffective for surface wave suppression. To address this issue, this study employed the proposed adaptive TFK filtering to process the two datasets. Figure 8 presents the filtering results of the two methods and their corresponding FK spectra, enabling an intuitive comparison of their surface-wave suppression performance and the associated spectral energy distribution characteristics.

For Dataset 1, Figure 8A shows the processing results of traditional FK filtering, while Figure 8B presents those of adaptive TFK filtering. For Dataset 2, Figure 8C,D displays the filtering results under the condition of mirror extension, where Figure 8C corresponds to traditional FK filtering and Figure 8D to adaptive TFK filtering. Boundary artifacts can be effectively mitigated by performing mirror extension and then extracting the region corresponding to the original dataset. By comparing Figure 8A with Figure 8B, and Figure 8C with Figure 8D, it can be observed that, compared with traditional FK filtering, adaptive TFK filtering achieved more significant surface wave suppression in both datasets. Meanwhile, it better preserved the continuity and integrity of reflection wave events, thus reducing damage to effective signals. A comparison between Figure 8B and Figure 8D demonstrated that adaptive TFK filtering exhibited superior surface-wave suppression performance to traditional FK filtering when mirror extension was applied. In contrast, traditional FK filtering suffered from obvious residual surface waves and energy dispersion, whereas the TFK method more effectively concentrated and suppressed surface wave energy.

The unilateral FK spectra in Figure 9 indicate significant energy overlap between surface and reflection waves in the FK domain for both field datasets, making it difficult for traditional FK filtering to achieve effective separation

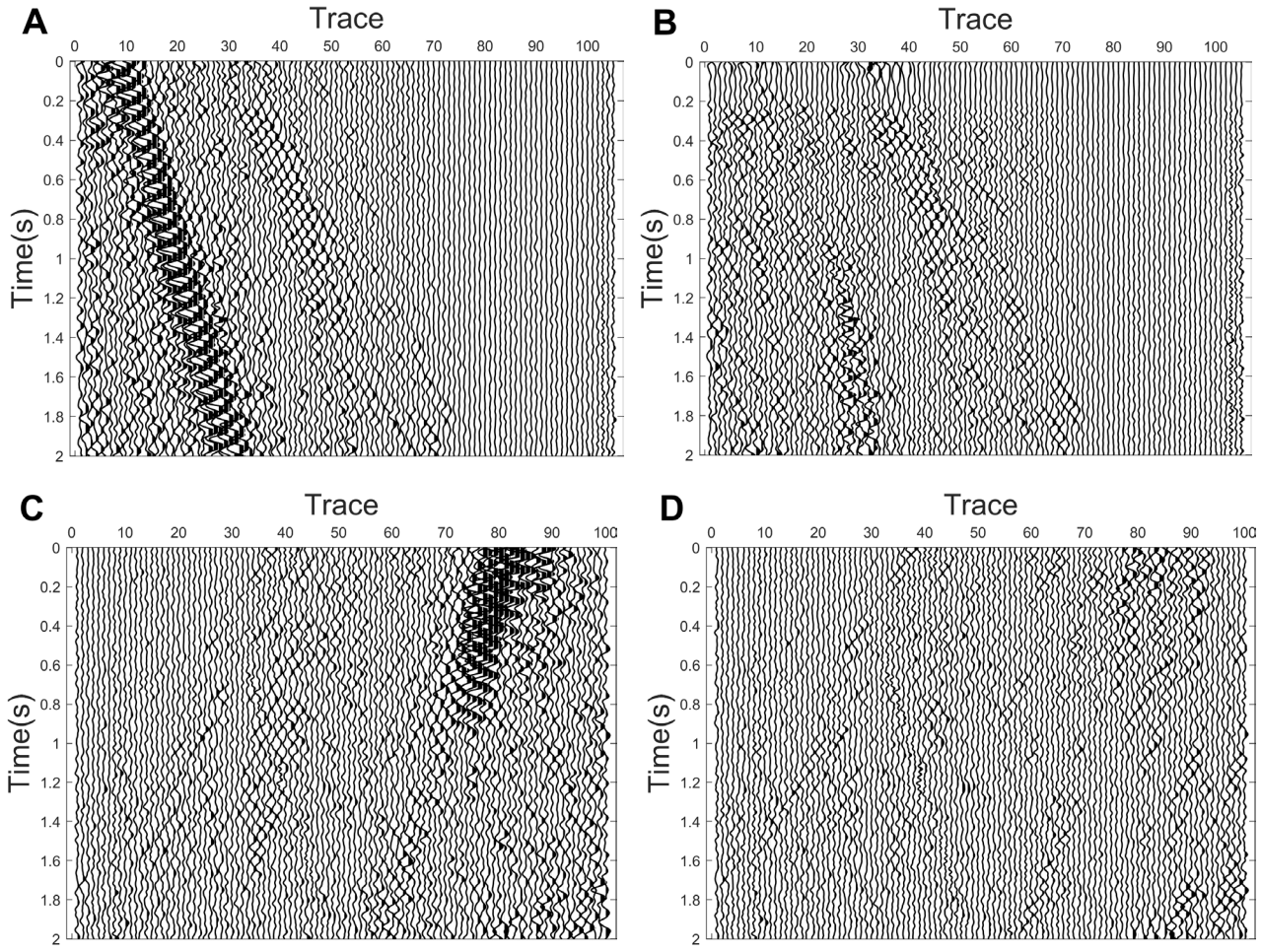



**Figure 8.** Surface-wave suppression results from the two field seismic datasets. (A) Frequency–wavenumber (FK) filtering on Dataset 1. (B) Adaptive time–frequency–wavenumber (TFK) filtering on Dataset 1. (C) FK filtering on Dataset 2. (D) Adaptive TFK filtering on Dataset 2.

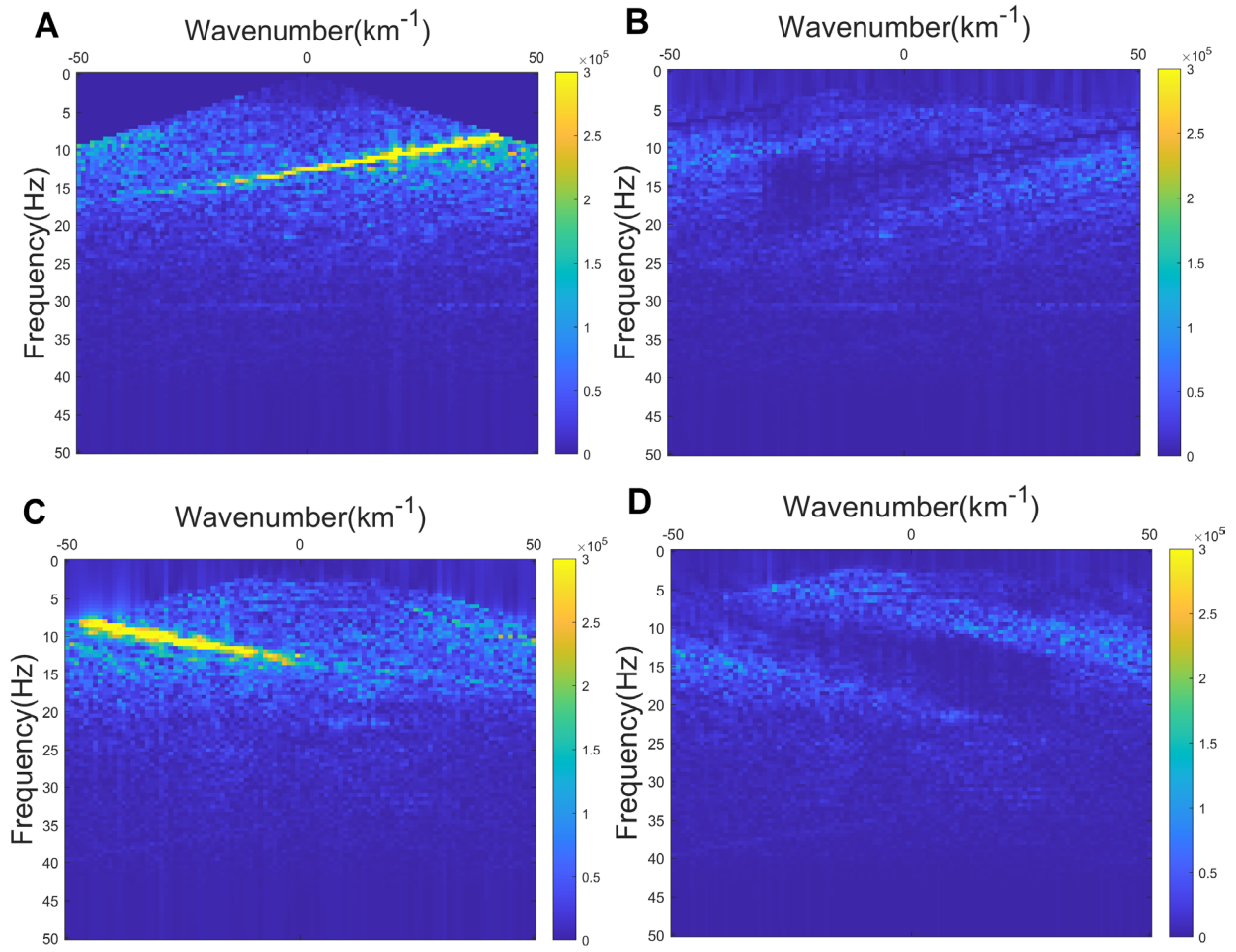



**Figure 9.** Comparison of FK filtering and adaptive time–frequency–wavenumber (TFK) filtering for the two field seismic datasets. (A) FK filtering on Dataset 1. (B) Adaptive TFK filtering on Dataset 1. (C) FK filtering on Dataset 2. (D) Adaptive TFK filtering on Dataset 2.

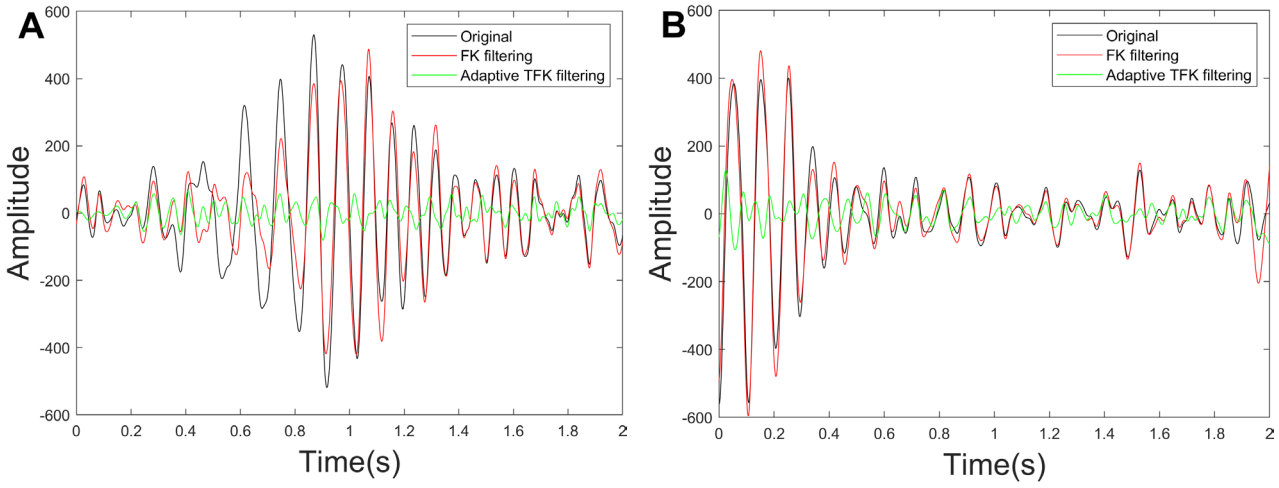



**Figure 10.** Comparison of frequency–wavenumber (FK) spectra after suppression in (A) the 20th trace from Dataset 1 and (B) the 80th trace from Dataset 2: original data (black line), FK filtering (red line), and adaptive time–frequency–wavenumber (TFK) filtering (green line)

without compromising reflection signals. By introducing the temporal dimension, adaptive TFK filtering improved the ability to characterize non-stationary energy distributions, thereby alleviating the aforementioned overlap issue to a certain extent and enabling more efficient surface wave suppression.

To provide a more precise trace-level comparison of filtering performance, representative traces were extracted from the two field datasets for detailed analysis (Figure 10): the 20th trace from Dataset 1 and the 80th trace from Dataset 2. For the 20th trace of Dataset 1 (Figure 10A), surface wave energy was mainly concentrated within 0.4–1.4 s, where the original record exhibited large amplitudes. Conventional FK filtering suppressed the surface waves only partially, leaving noticeable residuals. In contrast, the adaptive TFK filter yielded stronger suppression in this interval. Outside the surface wave window (0–0.4 s and 1.4–2 s), which was dominated by reflections, waveform changes were minor, indicating that the proposed method suppresses surface waves while preserving reflection events.

For the 80th trace of Dataset 2 (Figure 10B), surface wave energy was primarily distributed within 0–0.6 s. Conventional FK filtering again left residual surface wave energy, whereas the adaptive TFK filter provided further suppression. In the 0.6–2 s interval, the waveforms were largely preserved, suggesting limited distortion to effective signals. Overall, consistent improvements on representative traces from both datasets support the robustness and generality of the proposed approach.

To further evaluate the effectiveness of surface wave suppression, stacked sections from the unprocessed data, the conventional FK-filtered results, and the proposed adaptive TFK-filtered results were generated and compared (Figure 11). In the unprocessed data, strong surface waves appeared as high-amplitude, highly coherent noise bands,

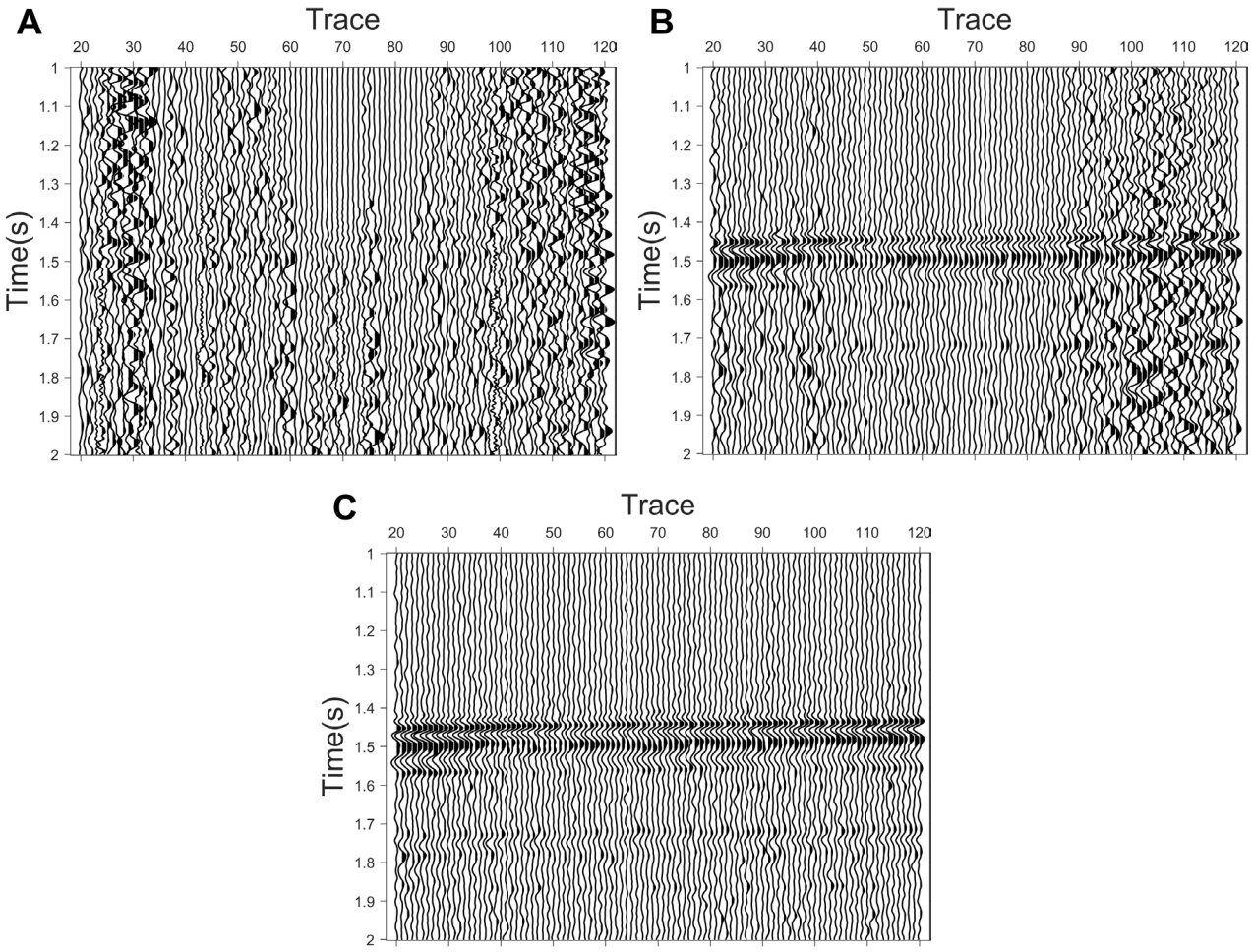



**Figure 11.** Profile comparison of surface wave suppression using different methods: (A) unprocessed data; (B) data after conventional frequency–wavenumber filtering; (C) data after the adaptive time–frequency–wavenumber filtering

severely obscuring reflection events and hindering reliable identification and tracking of reflection coherency. After applying the conventional FK filter, the primary reflection events became discernible; however, noticeable residual surface wave energy persisted, accompanied by locally degraded phase continuity. In contrast, the data obtained with the adaptive TFK filter showed further suppression of residual surface wave bands, with markedly improved continuity and traceability of reflection events. These observations indicate that the proposed method can more effectively suppress surface wave interference while largely preserving the dominant morphology and phase characteristics of the useful reflected signals.

In summary, the comparative results in Figures 8–11 demonstrate that, for field data in which surface wave and reflection energies overlap substantially in the FK domain, conventional FK filtering often fails to simultaneously achieve effective surface wave suppression and preserve reflection continuity. By incorporating the time dimension, the proposed adaptive TFK filtering leverages correlations between FK spectra from adjacent time windows to provide a data-driven characterization of surface wave-dominated regions. This enables a more effective spectral suppression of surface wave energy and its aliasing/mixing with reflection bands, while yielding reduced residual surface wave artifacts and more continuous, traceable reflection events at both the gather and stacked-section scales. These results validate the feasibility of targeted suppression in the TFK domain by exploiting differences in the time-varying spectral distributions of surface waves and reflections, and further suggest a more robust implementation pathway for adaptive surface wave suppression under complex field-acquisition conditions.

This study also conducted an efficiency comparison between the proposed method and FK filtering, focusing on processing efficiency under both simulated and field datasets. In principle, the TFK method operates via an FK transformation in the TFK domain, resulting in lower efficiency than conventional FK filtering. However, due to its parallelizable computational process, the method still holds substantial potential to improve efficiency in practical data processing. The computational time statistics are presented in Table 1.

**Table 1. Efficiency comparison**

| Sample | FK filter (s) | TFK filter (s) |
|---|---|---|
| Dataset 1 | 0.01 | 3.90 |
| Dataset 2 | 0.20 | 4.50 |

Abbreviations: FK: Frequency–wavenumber; TFK: Time–frequency–wavenumber.

As shown in Table 1, the TFK filtering method performed worse in single-core processing. However, in practical processing scenarios, it can still leverage GPU parallel computing, thereby alleviating this inefficiency to a certain extent.

## 4. Discussion

In this research, an adaptive TFK filtering method for surface wave suppression was proposed. The core idea of this approach is to construct an adaptive filter by cross-correlating spectral components in the FK domain across successive time intervals. This filter dynamically adjusts its characteristics in response to temporal variations in the seismic data, thereby better capturing the non-stationary nature of surface waves. After processing data in the FK domain, an inverse transform is applied to reconstruct the time-domain seismic records, preserving relative signal characteristics while effectively reducing surface wave energy. The key findings are as follows:

(i) Effectiveness in simulated data: The analysis of simulated datasets demonstrates that the proposed method can efficiently and automatically identify the spatial and spectral locations of surface waves within the FK domain. By adaptively adjusting filter parameters across different time windows, the method achieves enhanced discrimination between surface and effective waves. Compared with the conventional FK filtering technique, the proposed approach considerably simplifies the workflow by eliminating the need for repeated manual parameter tuning. Furthermore, it achieves rapid and accurate localization of surface waves, leading to more complete suppression of dispersive surface energy while maintaining the fidelity of the effective signal. Here, the term *fidelity* refers to the preservation of relative amplitude relationships of effective signals, rather than strict absolute amplitude preservation.

(ii) Improved stability and adaptability: Because the filter adapts dynamically with time, the method exhibits significantly improved stability and robustness. It can effectively accommodate temporal variations and changes in the statistical characteristics of seismic data, thereby ensuring more consistent performance across different datasets and acquisition conditions. This adaptability also reflects the intrinsic time-varying behavior of surface waves, implying that multi-scale and time-dependent information must be incorporated for efficient suppression. Consequently, the adaptive TFK method not only enhances the stability of the filtered data but also broadens the technique's applicability to a broader

range of geological settings and survey environments.

(iii) Superior performance in field data: When applied to real seismic field data, comparative analysis with traditional FK filtering further confirms the advantages of the proposed adaptive TFK approach. It demonstrates greater robustness in handling complex field conditions, effectively reducing surface wave interference while retaining weak yet interpretable effective signal components. The method requires fewer empirical adjustments and operates with acceptable computational efficiency for practical applications, while producing cleaner seismic records with improved signal-to-noise ratios. Overall, the adaptive TFK surface-wave suppression technique provides a practical, theoretically sound solution with acceptable processing efficiency for mitigating surface wave contamination in both simulated and field seismic datasets.

(iv) Limitations and future research directions: Notably, higher-mode surface waves were not removed in the present study. The core focus of this research is the suppression of high-energy surface waves, which are the primary source of interference affecting the quality of target reflected signals in seismic data processing. Within the datasets employed in this work (both simulated and field data), higher-mode surface waves exhibit relatively weak energy, exerting minimal impact on overall data interpretation. Therefore, effective extraction and suppression of higher-mode surface waves—particularly in scenarios where such modes are prominent—would require further investigation. Future efforts could focus on refining the adaptive identification logic in the TFK domain, tailoring key parameters to the energy and propagation features of low-energy higher-mode waves, or incorporating specialized modal separation mechanisms into the existing framework.

## 5. Conclusion

This study presents an adaptive TFK approach for surface wave suppression in seismic data. This method introduces a dynamic filtering framework that accounts for the time-varying and dispersive characteristics of surface waves, thereby achieving more accurate and adaptive noise suppression.

The findings indicate that the adaptive TFK method not only enhances surface wave suppression but also preserves the relative amplitude characteristics of effective signals. By constructing a time-dependent filtering strategy in the FK domain and integrating TF optimization, the approach demonstrates strong adaptability and stability across various seismic conditions. These improvements contribute to more reliable data for subsequent imaging and inversion tasks.

The proposed technique provides a practical and efficient alternative to conventional FK filtering, reducing the need for manual parameter adjustments and offering greater robustness in complex field-acquisition environments. Its flexibility and computational efficiency make it a promising tool for modern seismic data processing workflows.

## Acknowledgments

None.

## Funding

This research was supported by the Open Fund of the SINOPEC Key Laboratory of Geophysics (Grant No. 36750000-24-FW0399-0004), the Natural Science Foundation of Xinjiang Uygur Autonomous Region (Grant No. 2022D01C659) under the "Tianchi Talent" Introduction Program, and the Open Fund of Northwest Oilfield Branch Company, SINOPEC (Grant No. 34400000-25-ZC0607-0156).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Yong Wang
*Formal analysis:* Yong Wang, Xuejie Gao
*Investigation:* Xuejie Gao, Zerun Nian, Lang Yang
*Methodology:* Yong Wang, Youjuan He, Zhili Chen
*Writing–original draft:* Xuejie Gao
*Writing–review & editing:* Youjuan He

## Availability of data

Data associated with this research are available and can be obtained by contacting the corresponding author.

## References

1. Zheng JJ, Yin XY, Zhang GZ, Wu GH, Zhang ZS. The surface wave suppression using the second generation curvelet transform. *Appl Geophys*. 2010;7(4):325–335.

   doi: 10.1007/s11770-010-0257-x

2. Wang D, Li Z, Huang J, Dong L, Ding C. Surface wave suppression based on the characteristic of reflected wave. In: *Proceedings of Beijing 2014 International Geophysical Conference & Exposition*; 21-24 April 2014; Beijing, China. Society of Exploration Geophysicists and Chinese Petroleum Society; 2014:307-310.

   doi: 10.1190/igcbeijing2014-079

3. Lv S, Gong X, Hu B, Xu Z, Zhang J, Cheng Q. Adaptive non-subsampled shearlet transform and its application to surface wave suppression. *IEEE Trans Geosci Remote Sens*. 2024;62:1–15.

doi: 10.1109/TGRS.2024.3477954

4. Li Q, Shen A, Guo Y, Wu J, Shi Y. Review of evaluating asphalt pavement structure integrity and strength with Rayleigh wave methods: techniques, applications, and trends. *Measurement*. 2024;234:114849.

doi: 10.1016/j.measurement.2024.114849

5. Tokeshi J, Karkee CK, Sugimura Y. Reliability of Rayleigh wave dispersion curve obtained from f–k spectral analysis of microtremor array measurement. *Soil Dyn Earthq Eng*. 2006;26(2–4):163–174.

doi: 10.1016/j.soildyn.2005.02.013

6. Kong XL, Chen H, Hu ZQ, Kang JX, Xu TJ, Li LM. Surface wave attenuation based on polarization attributes in time-frequency domain for multicomponent seismic data. *Appl Geophys*. 2018;15(1):99–110.

doi: 10.1007/s11770-018-0656-y

7. Rasol MA, Pérez-Gracia V, Fernandes FM, Pais JC, Santos-Assunçao S, Santos C, Sossa VG. GPR laboratory tests and numerical models to characterize cracks in cement concrete specimens exemplifying damage in rigid pavement. *Measurement*. 2020;158:107662.

doi: 10.1016/j.measurement.2020.107662

8. Meng A, Xu H, Feng X, Tan Y. Feasibility of freeze-thaw damage analysis for asphalt mixtures through dynamic nondestructive testing. *Constr Build Mater*. 2020;233:117220.

doi: 10.1016/j.conbuildmat.2019.117220

9. Wang X, Huang H, Shen S, Jin G, Mao Q, Lu H. Characterization of in situ modulus of asphalt pavement and its relation to cracking performance using SASW method. *J Transp Eng Part B Pavements*. 2020;146(3):04020039.

doi: 10.1061/JPEODX.0000195

10. Yuanyuan H, Rui Y, Jishun P, Xin L. A 3D surface wave attenuation method based on dispersion curve analysis and its application. *Acta Geophys*. 2024;72(6):4123–4138.

doi: 10.1007/s11600-024-01318-2

11. Serdyukov AS, Yablokov AV, Duchkov AA, Azarov AA, Baranov VD. Slant f–k transform of multichannel seismic surface wave data. *Geophysics*. 2019;84(1):A19–A24.

doi: 10.1190/geo2018-0430.1

12. Serdyukov AS. Ground-roll extraction using the Karhunen–Loeve transform in the time-frequency domain. *Geophysics*. 2022;87(2):A19–A24.

doi: 10.1190/geo2021-0453.1

13. Zheng Y, He DK, Feng X. Surface wave suppressing and decomposition using deep neural networks. In: Proceedings of 83rd EAGE Annual Conference & Exhibition; 6–9 Jun 2022; Madrid, Spain. European Association of Geoscientists & Engineers; 2022:1-5.

doi: 10.3997/2214-4609.202210218

14. Jin Z, Li G, Ou G. Improvement of inverse dispersion technique for ground roll attenuation in sedimentary environment. *J Appl Geophys*. 2024;220:105268.

doi: 10.1016/j.jappgeo.2023.105268

15. Liu S, Birnie C, Bakulin A, Dawood A, Silvestrov I, Alkhalifah T. A self-supervised scheme for ground roll suppression. *Geophys Prospect*. 2024;72(7):2580–2598.

doi: 10.1111/1365-2478.13522

16. Son YH, Park H, Cho Y, Min DJ. Ground-roll attenuation using dual-model self-supervised selective learning with blind horizontal convolutional neural networks. *J Appl Geophys*. 2024;224:105363.

doi: 10.1016/j.jappgeo.2024.105363

17. Yablokov AV, Moiseev MV, Serdyukov AS, Litvichenko DA. Algorithm for surface wave suppression on 2D seismic data using the slant Karhunen–Loeve transform in a time-frequency domain. *Russ Geol Geophys*. 2024;65(5):650–662.

doi: 10.2113/RGG20234660

18. He XL, Liu SQ, Tong ZQ. 3-D denoising method through visual speed in frequency–wavenumber domain. *J China Univ Pet (Nat Sci Ed)*. 2010(4):62–66. [In Chinese].

doi: 10.3969/j.issn.1673-5005.2010.04.011

19. Tian Y, Wang X, Peng G, Wang X, Su Q. Noise attenuation technology on wide azimuth seismic data. *Oil Geophys Prospect*. 2013;48(2):187–191. [In Chinese].

doi: 10.13810/j.cnki.issn.1000-7210.2013.02.013

20. Askari R, Siahkoohi HR. Ground roll attenuation using the S and x–f–k transforms. *Geophys Prospect*. 2008;56(1):105–114.

doi: 10.1111/j.1365-2478.2007.00659.x

21. Liu Y, Fomel S. Seismic data analysis using local time–frequency decomposition. *Geophys Prospect*. 2013;61(3):516–525.

doi: 10.1111/j.1365-2478.2012.01062.x

22. Wang C, Wang Y. Ground roll attenuation using polarization analysis in the t–f–k domain. *Geophys J Int*. 2017;210(1):240–254.

doi: 10.1093/gji/ggx152

23. Wu Y, Pan S, Yi S, Chen J, Cui Q, Song G. An automatic screening method for the passive surface-wave imaging based on the F–K domain energy characteristics. *IEEE Trans Geosci Remote Sens*. 2023;61:1–12.

doi: 10.1109/TGRS.2023.3321786

24. Sun H, Liu E, Zhang F, Wang W, Li X. Seismic shear wave noise suppression and application to well tie. *J Geophys Eng*. 2024;21(6):1619–1635.

doi: 10.1093/jge/gxae091

25. Stockwell RG, Mansinha L, Lowe RP. Localization of the complex spectrum: the S transform. *IEEE Trans Signal Process*. 1996;44(4):998–1001.

doi: 10.1109/78.492555

26. Xu S, Zhang Y, Lambaré G. Antileakage Fourier transform for seismic data regularization in higher dimensions. *Geophysics*. 2010;75(6):WB113–WB120.

doi: 10.1190/1.3507248.

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Reducing interpretation risks in organic-rich low-velocity mudstone using rock physics modeling in Yinggehai Basin, South China Sea

**Yanhua Xie, Jinpeng Li*, Fang Li, Wanyuan Sun, and Yayun Dang**

Department of Geophysics, Research Institute, China National Offshore Oil Corporation Limited Hainan Branch, Haikou, Hainan, China

(This article belongs to the *Special Issue: Geophysical Inversion and Intelligent Prediction Technologies for Complex Hydrocarbon Reservoirs*)

## Abstract

The mudstone in a specific exploration area of the Yinggehai Basin is rich in organic matter and serves as a paradigmatic example of a low-velocity mudstone formation in rock physics. This type of interval exhibits seismic response characteristics similar to those of hydrocarbon reservoirs, which complicates the identification of gas reservoirs and the prediction of gas-bearing zones in this area. In this study, the analysis and modeling of rock physical characteristics within the exploration area are investigated. Based on the actual drilling curve and coring analysis data, a rock physics model of the low-velocity mudstone reservoir is established using self-consistent approximations and the Ciz–Gassmann model, and the influence of organic matter content on the mudstone's elastic properties is analyzed. Additionally, a robust quantitative inversion procedure is introduced to test the feasibility of inverting for porosity, clay content, and water saturation to mitigate the risk of low-velocity mudstone. The application of actual data, including core samples, logging data, and seismic database, demonstrates the effectiveness of this method and provides technical support for seismic prediction of sand body identification and hydrocarbon detection.

***Keywords:*** Low-velocity mudstone; Rock physics; Hydrocarbon detection; Total organic carbon; Seismic interpretation

***Corresponding author:**
Jinpeng Li
(lijp40@cnooc.com.cn)

**Citation:** Xie Y, Li J, Li F, Sun W, Dang Y. Reducing interpretation risks in organic-rich low-velocity mudstone using rock physics modeling in Yinggehai Basin, South China Sea. *J Seismic Explor.* 2026;35(2):025460110. doi: 10.36922/JSE025460110

**Received:** November 14, 2025

**Revised:** February 2, 2026

**Accepted:** February 5, 2026

**Published online:** April 30, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

The Yinggehai Basin is located in the western part of the northern continental shelf of the South China Sea. Some gas reservoirs exhibit typical bright-spot seismic response characteristics of low frequency and strong amplitude[1–5], which is beneficial for target search. As exploration and development have progressed, the error rate of target search using strong amplitude characteristics in the Yinggehai Basin has gradually increased.[6–8] It can thus be concluded that the atypical reflection characteristics exhibited by low-frequency, strong-amplitude bright spots prior to drilling are not indicative of a response from gas-bearing sandstone (GS). Rather, these phenomena are frequently the result of false bright spots, which are often caused by lithology changes. A significant and archetypal

rationale pertains to the impact of low-velocity mudstone (LVM). There is no strict correspondence between proven gas reservoirs and strong seismic reflection characteristics. The issue of how to avoid LVM traps has become a pressing concern in exploration practice.[9,10]

Petrophysics bridges physical property parameters to elastic responses. The inverse problem in petrophysical modeling involves predicting reservoir parameters based on petrophysical models. The fundamental principle of this process involves the extraction of reservoir parameters from pre-stack seismic or elastic data. This process decouples different lithologic and physical property layers.[11] For instance, incorporating rock physics statistical patterns and prior information on physical properties into seismic forward-modeling objective functions enables inversion for porosity and impedance parameters.[12–14] Alternatively, establishing mappings between physical parameters and saturated rock velocities using equivalent medium theory and wave-induced fluid flow seismic wave dispersion equations enables the solution of inverted parameters within Bayesian pre-stack amplitude-variation-with-offset (AVO) inversion or elastic impedance inversion frameworks.[15,16] The existing body of research indicates that reservoir parameter extraction based on rock physics modeling is a reliable and effective method for reservoir prediction and hydrocarbon detection.[17,18]

Targeted research on petrophysical modeling and seismic prediction methods is essential to analyze the seismic petrophysical characteristics of LVM, clarify their petrophysical properties, mitigate associated risks, and enhance reservoir prediction and hydrocarbon detection success rates. The present study analyzes and summarizes the petrophysical characteristics of LVM encountered in a drilling area of the Yinggehai Basin, based on actual data from the western South China Sea exploration zone. For typical LVM formations with rich-type mechanisms, rock-physical modeling and analysis are conducted based on the equivalent-medium theory. The velocity inversion prediction is achieved using actual well data. The establishment of a quantitative seismic prediction methodology tailored for formations containing LVM represents an advancement in the field. This methodology provides a sound theoretical and methodological basis for reservoir prediction and hydrocarbon detection.

## 2. Geological setting

The Yinggehai Basin is located in the sea area to the east of the Indo–China Peninsula and to the west of Hainan Island. It includes first-order tectonic units, such as the central depression, Yingdong slope, and Yingxi slope.[1,2] The basin is characterized by Cenozoic overpressure and high temperature, and is considered to be a significant area for the exploration and development of natural gas resources in the South China Sea. Numerous large and medium-sized gas fields have been discovered and brought into production in the Yinggehai Basin. The location of the target area in the Yinggehai Basin is shown in Figure 1. The formation, environment, and petroleum source of this basin are shown in Figure 2. The sedimentary evolution of the Yinggehai Basin from the late Oligocene Lingshui Formation to the Miocene Yinggehai Formation has been shown to record a transition from late rifting to a depression and an accelerated subsidence period. The Lingshui Formation is deposited in the sea–land transition zone in a shallow-sea environment and developed multi-type reservoirs, including delta, shore-shallow sea sandstone, and carbonate platform reservoirs. The provenance is evident from northwest, west, and other directions, and the characteristics of sand–shale interbeds are clearly discernible.[3] The Yinggehai Formation has transitioned from a deep-sea to a semi-deep-sea environment, with the Red River as the dominant source. The formation of thick mudstones is widespread, and isolated deep-water turbidite sandstones are present, forming a natural gas accumulation system under high-temperature, high-pressure conditions. The presence of mud diapirs is identified as the primary control factor within this system. The evolution of this sedimentary sequence indicates the transformation of the basin from a rift-divided lake basin to a unified, rapidly subsiding basin. The sedimentary system is characterized by the predominance of gravity flow over traction flow. The reservoir gradually transformed from a horizontal, continuous layered reservoir to a lenticular lithologic body. This unique oil and gas geological model comprises a lower generation and an upper reservoir, with high temperatures and pressures, and diapir control within the basin.[7,8]

## 3. Data and methods

### 3.1. Data

Low-velocity mudstone is a geophysical relative concept, signifying that the P-wave velocity ($V_p$) of this particular mudstone is lower in comparison to that of its adjacent normal mudstone or sandy mudstone strata.[2,3] A significant decrease in velocity leads to the formation of a strong wave impedance interface between this type of mudstone and the high-speed formation below, and this characteristic will result in strong reflection amplitude on the seismic profile, which is a typical seismic characteristic of gas-bearing reservoirs. While performing geosteering during drilling operations, if strong seismic reflection amplitudes are the sole criterion for judgment, there is a risk of misidentifying the LVM layer as a valuable GS layer.[4,5] This can deviate the

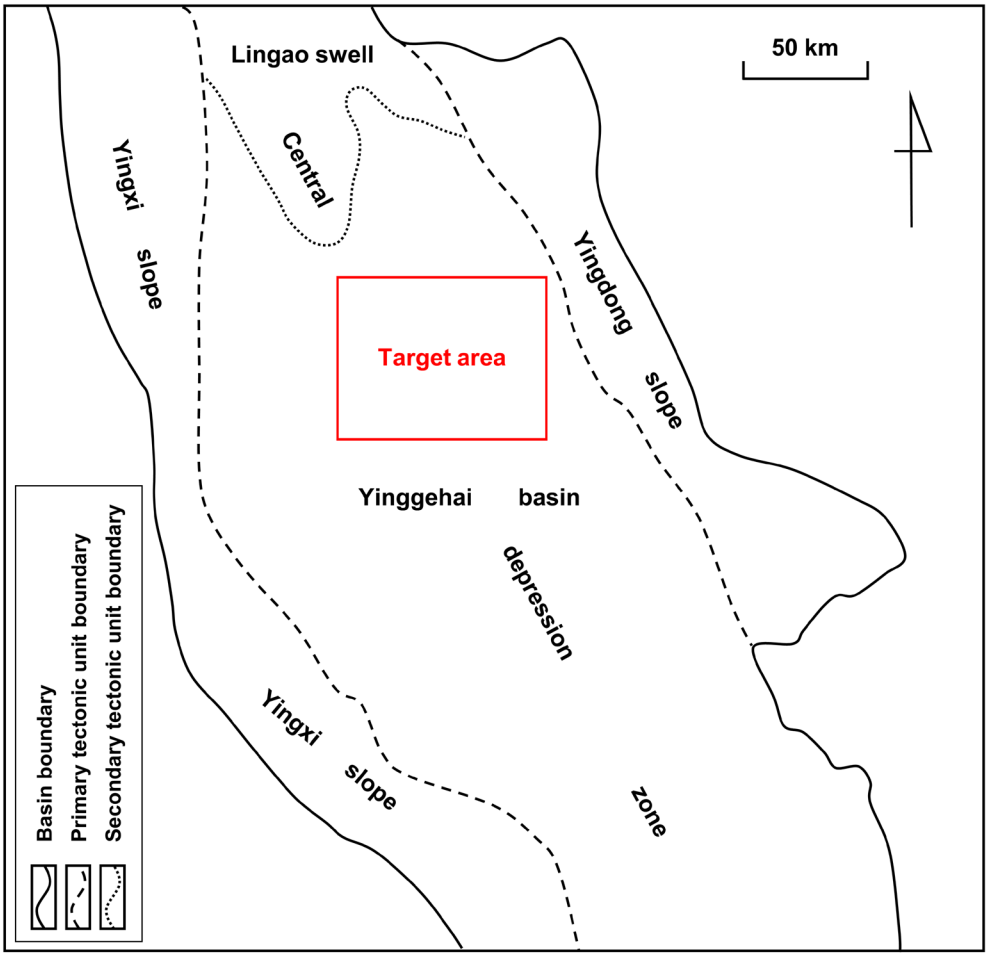



**Figure 1**. Location of the target area in the Yinggehai Basin

drilling trajectory from the actual target reservoir, leading to invalid footage and economic losses.[6]

Exploration practices suggest that the mechanisms that induce low-velocity characteristics in the region's mudstones include three factors: pore pressure (p) anomalies, effects of clay minerals, and organic matter content. Statistical data are presented in Table 1 and Figure 3. The first category of LVM exhibited primarily low resistivity ($R_t$), a high P- and S-wave velocity ratio ($V_p/V_s$), and a high pore pressure coefficient ($p$ coefficient). However, its gamma ray (GR), total organic carbon (TOC), and gas-detected total hydrocarbon values were comparable to those of high-velocity background mudstone (BM). TOC content can represent the abundance of organic matter in rocks. This LVM's internal composition differs little from that of normal-velocity BM. The main distinction lies in the external environment, specifically a significantly higher $p$ coefficient than that of the BM. The second category of LVM exhibited similar characteristics to the first category in terms of $R_t$, $V_p/V_s$, TOC, gas-detected total hydrocarbons, and $R_t$ versus $V_p$. However, it showed significant differences in gamma curves, with GR values markedly higher than those of the normal-velocity mudstone background, while the $p$ coefficient is comparable. This indicates that the pressure environment of this LVM is not substantially different from that of normal-velocity mudstone backgrounds, with the primary distinction being clay mineral content, which was markedly higher than that of normal-velocity mudstone backgrounds. The third category of LVM differed from the first and second categories. Although differences in GR, $R_t$, and the $p$ coefficient were not pronounced compared to the normal-velocity BM, the third category of LVM showed markedly lower $V_p/V_s$ and higher TOC, overall resembling the characteristics of gas-bearing strata. This LVM category showed minimal differences from BM in external temperature and pressure conditions and clay content, with the primary distinction in organic carbon content.

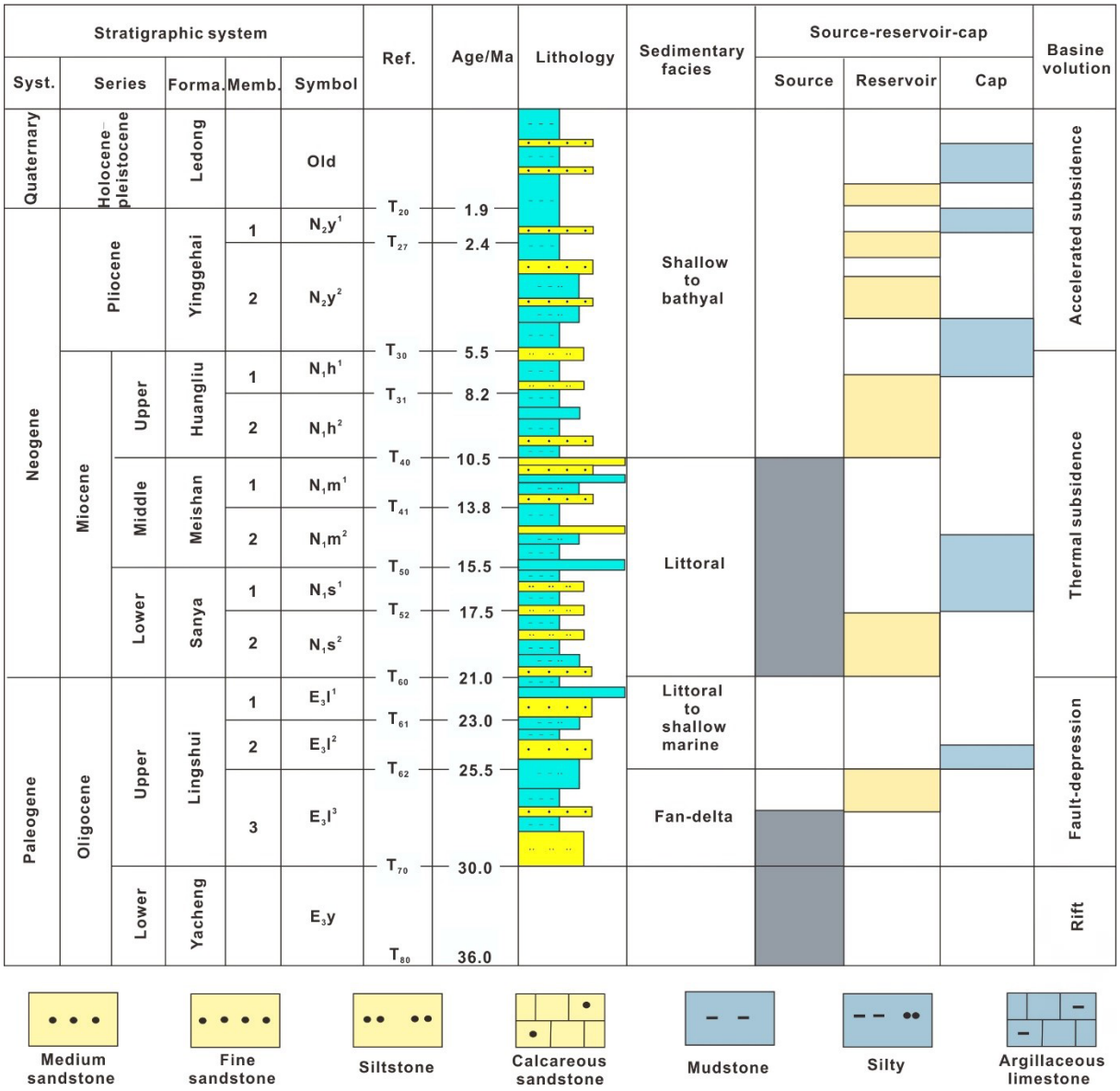



**Figure 2.** Formation, environment, and the petroleum source of the Yinggehai Basin

**Table 1. Statistical table of geophysical characteristics of low-velocity mudstone in Yinggehai Basin**

| Layers | Parameter characteristics | | | | | | Cause of low-velocity |
|---|---|---|---|---|---|---|---|
| | Resistivity | GR | $V_p/V_s$ | TOC | Pore pressure coefficient | Gas detection, total hydrocarbon | |
| Category 1 | Low | Similar | High | No | High | Similar | High pressure |
| Category 2 | Low | High | High | No | Similar | Similar | Massive clay |
| Category 3 | High/similar | High/similar | Low/similar | Yes | Similar | High | High TOC |

Abbreviations: GR: Gamma ray; TOC: Total organic carbon; $V_p/V_s$: P- and S-wave velocity ratio.

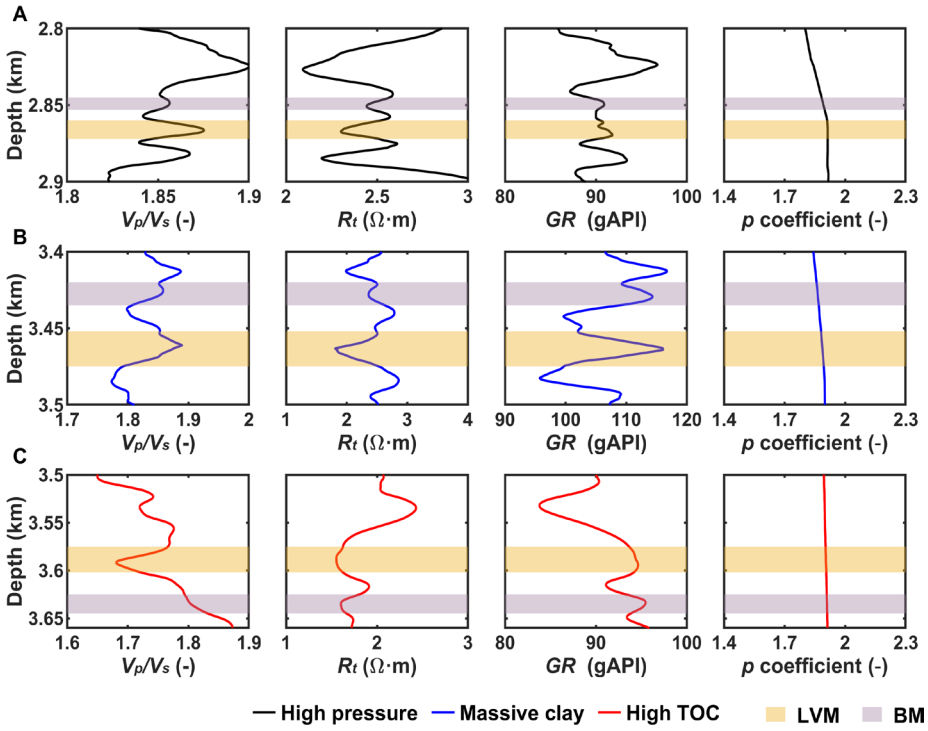



**Figure 3.** P- and S-wave velocity ratio ($V_p/V_s$), resistivity ($R_t$), gamma ray (GR), and pore pressure ($p$) coefficient that incorporates the data in Table 1. (A) First category of LVM. (B) Second category of LVM. (C) Third category of LVM.
Abbreviations: BM: Background mudstones; LVM: Low-velocity mudstone; TOC: Total organic carbon.

Figure 4A illustrates the correlation between $V_p$ and clay content in Well A, revealing two distinct trends in mudstone velocity at varying clay contents. Notably, LVM exhibits significantly lower $V_p$ than BM at similar clay contents. In addition, red and blue solid scatters represent GS and water-bearing sandstone, respectively. Figure 4B shows the relationship between mudstone $V_p$ and $p$. It indicates that, as the $p$ coefficient increases, mudstone $V_p$ decreases significantly.

Similarly, Figure 5A illustrates the correlation between $V_p$ and clay content in Well B. While the LVM with high clay content exhibited a trend consistent with the BM, the LVM in this layer exhibited a significantly lower value $V_p$ due to its higher clay content. This low velocity may be related to the higher clay content, whereby clay functions as a skeletal component and reduces $V_p$.[6,10] For LVM with high TOC, however, the relationship between velocity and clay content exhibited a distinct trend compared to that of high clay content. Their velocities were markedly lower than those of BMs with equivalent clay content. However, as illustrated in Figure 5B, the relationship between $V_p$ and organic carbon content shows a clear decrease in velocity with increasing organic carbon content. The low velocity characteristics of high organic carbon are primarily induced by TOC enrichment. Conventional seismic inversion methods struggle to distinguish this type of mudstone from GS reservoirs.

Compared to clay minerals, organic matter (primarily kerogen) has physical properties characterized by a low bulk modulus, a low shear modulus, and a low density.[19] The target exploration area's reservoir minerals are predominantly quartz and clay. Geochemical analysis indicated that the organic matter consists primarily of immature kerogen. Mudstone has low porosity and permeability, with pore fluids consisting mainly of water and gas.

### 3.2. Key steps of rock physics modeling

Figure 6 illustrates the rock physics modeling workflow, accounting for the physical properties of kerogen. First, the elastic modulus of the sand–clay matrix was calculated using the Voigt–Reuss–Hill (VRH) averaging principle. Meanwhile, the rock–skeleton modulus for pores containing TOC was computed using the self-consistent approximations (SCA) theory. To ensure the low-frequency assumption, the SCA theory was not directly applied for TOC filling. Instead, the Ciz–Gassmann equation was used for solid replacement, incorporating TOC as a pore-filler

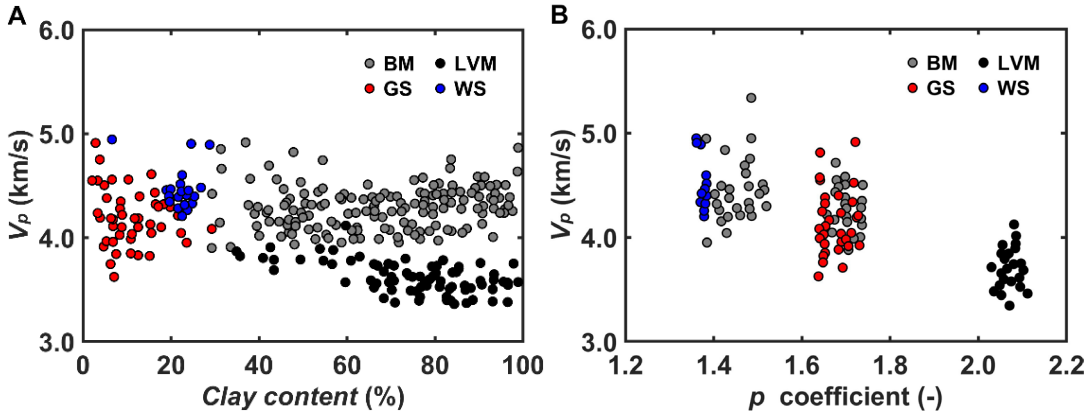



**Figure 4.** The relationship between (A) $V_p$ and clay content, and (B) $V_p$ and pore pressure coefficient of Well A
Abbreviations: BM: Background mudstone; GS: Gas-bearing sandstone; LVM: Low-velocity mudstone; WS: Water-bearing sandstone.

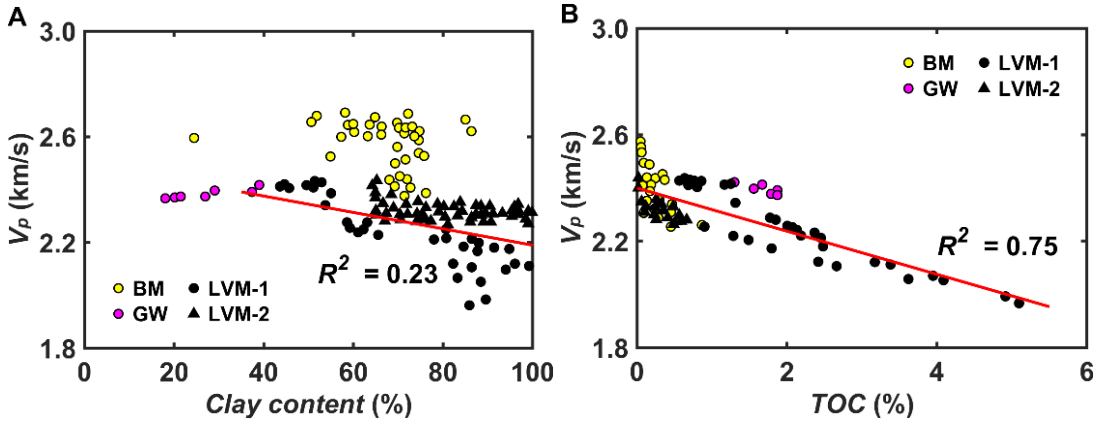



**Figure 5.** The relationship between (A) $V_p$ and clay content, and (B) $V_p$ and TOC of Well B
Abbreviations: BM: Background mudstone; GW: Gas-water-bearing sandstone; LVM: Low-velocity mudstone; TOC: Total organic carbon.

into the rock matrix to create a new equivalent matrix. Finally, SCA theory and the Gassmann equation were applied to calculate the elastic modulus of the equivalent matrix and the saturated rock. This ultimately yielded the $V_p$ and $V_s$.

The elastic properties of saturated rocks are directly related to the rock matrix, rock skeleton, and fluid filling, and exhibit a nonlinear coupled relationship.[20–22] Seismic petrophysical modeling of LVM containing TOC involves five key steps that incorporate five fundamental petrophysical models: (i) the VRH average theory, (ii) the SCA model, (iii) the Ciz–Gassmann equation, (iv) the pore fluid modulus model, and (v) the Biot–Gassmann fluid replacement equation.

### 3.2.1. Rock matrix modulus

The elastic properties of rock are primarily contingent on the composition of the rock matrix, with the type and amount of diagenetic minerals present playing a pivotal role in determining the rock's specific elastic characteristics. In the field of rock physics, the VRH average model is typically employed to estimate the equivalent modulus of the rock matrix, which can be expressed as **Equation 1**[23]:

$$M_{\mathrm{m}} = \frac{1}{2}\left[\sum_{i=1}^{N} f_i M_i + \left(\sum_{i=1}^{N} f_i / M_i\right)^{-1}\right] \quad (1)$$

In the formula, $M_{\mathrm{m}}$ is defined as the equivalent bulk modulus, or shear modulus ($K_{\mathrm{m}}$, $G_{\mathrm{m}}$), of the matrix. $M_i$ is expressed as the bulk modulus, or shear modulus ($K_i$, $G_i$), of the mineral components. $f_i$ is represented as the content of the mineral components, and $N$ is designated as the number of types of mineral components. The focus of this study is the exploration area, which primarily encompasses quartz and clay minerals, with $N = 2$.

### 3.2.2. Dry rock skeleton modulus

The porosity and pore structure of rock are significant factors in determining its elastic properties. The central theme of rock physics research is the establishment of a reasonable and equivalent description of the relationship

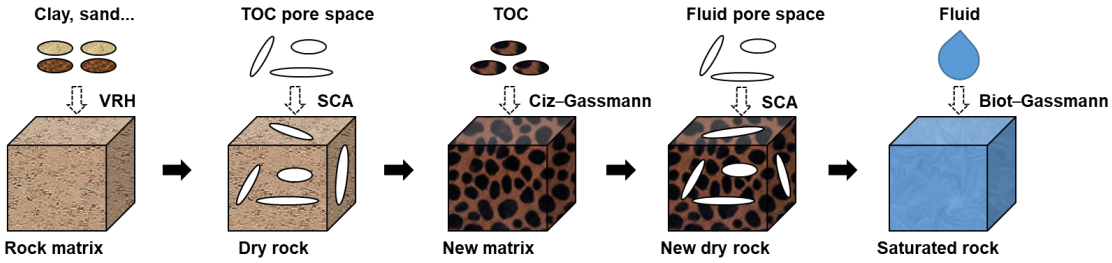



**Figure 6.** Rock physical modeling process of low-velocity mudstone containing TOC
Abbreviations: SCA: Self-consistent approximation; TOC: Total organic carbon; VRH: Voigt–Reuss–Hill.

between pore parameters and elastic parameters. In the context of the SCA theory, the porosity of the rock is equivalent to that of ellipsoids with a specific aspect ratio. Furthermore, the rock is equivalent to an equivalent medium composed of a matrix and inclusions with a particular aspect ratio. Therefore, the equivalent modulus of the rock skeleton can be calculated using the established method. The formula is expressed as **Equations 2** and **3**[24,25]:

$$\sum_{i=1}^{n} v_i \left( K_i - K_{\mathrm{SC}}^{*} \right) P^{*i} = 0 \tag{2}$$

$$\sum_{i=1}^{n} v_i \left( G_i - G_{\mathrm{SC}}^{*} \right) Q^{*i} = 0 \tag{3}$$

where $n$ denotes the total number of inclusion types, $v_i$ is the integral number of inclusion objects, and $K_{\mathrm{SC}}^{*}$ and $G_{\mathrm{SC}}^{*}$ are the equivalent bulk modulus and shear modulus, respectively. The coefficients $P^{mi}$ and $Q^{mi}$ are known as the polarization factors, defined as the effects of the $i$-th inclusion on the background matrix, respectively. $K_i$ and $G_i$ are the bulk modulus and shear modulus of the filler, respectively.

### 3.2.3. Solid substitution equations

In the case of the organic matter medium, which is characterized by its viscous properties, the solid substitution equations were employed to calculate the bulk modulus, denoted by $K_{\mathrm{sat}}$, and the shear modulus, denoted by $G_{\mathrm{sat}}$, when the pore space contains solids (**Equations 4** and **5**).[26]

$$K_{\mathrm{sat}}^{-1} = K_{\mathrm{dry}}^{-1} - \frac{\left( K_{\mathrm{dry}}^{-1} - K_{\mathrm{gr}}^{-1} \right)^2}{\phi \left( K_{if}^{-1} - K_{\phi}^{-1} \right) + \left( K_{\mathrm{dry}}^{-1} - K_{\mathrm{gr}}^{-1} \right)} \tag{4}$$

$$G_{\mathrm{sat}}^{-1} = G_{\mathrm{dry}}^{-1} - \frac{\left( G_{\mathrm{dry}}^{-1} - G_{\mathrm{gr}}^{-1} \right)^2}{\phi \left( G_{if}^{-1} - G_{\phi}^{-1} \right) + \left( G_{\mathrm{dry}}^{-1} - G_{\mathrm{gr}}^{-1} \right)} \tag{5}$$

It is imperative to note that $K_{\mathrm{dry}}$ and $G_{\mathrm{dry}}$ represent the bulk modulus and shear modulus of the dry rock skeleton, respectively, whereas $K_{\mathrm{gr}}$ and $G_{\mathrm{gr}}$ denote the bulk modulus and shear modulus of the background matrix, respectively. $K_{if}$ and $G_{if}$ are defined as the bulk modulus and shear modulus, respectively, with respect to the pore filler. $K_{\phi}$ and $G_{\phi}$ represent the pore-related bulk modulus and shear modulus, respectively. In accordance with the isotropic assumption, it can be demonstrated that $K_{\phi} = K_{\mathrm{gr}}$ and $G_{\phi} = G_{\mathrm{gr}}$.

### 3.2.4. Pore fluid modulus

The reservoirs within the exploration area consist of low-permeability clastic rocks, situated in the middle and deep layers. The pore structure is complex, and the gas–water immiscible fluid generally does not exhibit a highly relaxed, uniform filling state. A reasonable situation is that one fluid is filled into another fluid background in the form of patches, called the patch saturation state. The Brie model can be used to estimate the modulus of the mixed fluid (**Equation 6**)[27]:

$$K_{\mathrm{f}} = (K_{\mathrm{w}} - K_{\mathrm{hc}})(1 - S_{\mathrm{hc}})^{e} + K_{\mathrm{hc}} \tag{6}$$

where $K_{\mathrm{w}}$ and $K_{\mathrm{hc}}$ represent the bulk modulus of water and hydrocarbons, respectively. Furthermore, $S_{\mathrm{hc}}$ denotes hydrocarbon saturation, whilst $e$ is an empirical constant that is contingent on the specific data under consideration.

### 3.2.5. Biot–Gassmann model

The medium under consideration is a fluid–solid two-phase medium. The Gassmann equation establishes the relationship among the rock matrix, fluid, and pore. To study the influence of fluid on the elastic properties of rock in the seismic frequency band, it is generally necessary to use the Gassmann equation to fill or replace the fluid. It has been demonstrated that the prediction result of the Gassmann equation is equivalent to the low-frequency limit of the Biot theory, and this result is known as the Biot–Gassmann model (**Equations 7** and **8**)[28]:

$$\tilde{K}_{\mathrm{sat}} = \tilde{K}_{\mathrm{dry}} + \alpha^2 \left[ (\alpha - \phi)/K_{\mathrm{s}} + \phi/K_{\mathrm{f}} \right]^{-1} \tag{7}$$

$$\tilde{G}_{\text{sat}} = \tilde{G}_{\text{dry}} \tag{8}$$

where $\tilde{K}_{\text{sat}}$ and $\tilde{G}_{\text{sat}}$ are the bulk modulus and shear modulus of saturated rock, respectively. $\tilde{K}_{\text{dry}}$ and $\tilde{G}_{\text{dry}}$ represent the skeleton modulus of rocks filled with TOC, respectively. $\alpha$ is the Biot coefficient and $\alpha = 1 - \tilde{K}_{\text{dry}} / K_{\text{m}}$.

Following the acquisition of the modulus and density of saturated rock, the $V_p$ and $V_s$ of the rock medium can be calculated as in **Equations 9** and **10**:

$$V_p = \sqrt{\frac{\tilde{K}_{\text{sat}} + 4\tilde{G}_{\text{sat}}/3}{\rho_{\text{sat}}}} \tag{9}$$

$$V_s = \sqrt{\frac{\tilde{G}_{\text{sat}}}{\rho_{\text{sat}}}} \tag{10}$$

where $\rho_{\text{sat}}$ denotes the saturated rock density.

### 3.3. Numerical simulation

To verify the rationality and applicability of the rock physics model developed in this study and to clarify the influence of parameters, such as organic matter content, clay content, and porosity, on the reservoir's elastic properties, a sand–mudstone forward simulation model was developed. The model under consideration was a sand–clay reservoir with TOC, with gas and water as the pore fluids. The TOC ranged from 1% to 6%, the porosity varied from 5% to 15%, and the clay content varied from 10% to 70%. As shown in Table 2, the fundamental parameters involved in the numerical example were derived from previous studies.[29–31]

**Table 2. Numerical simulation parameters[29–31]**

| Parameters | Bulk modulus (GPa) | Shear modulus (GPa) | Density (g·cm$^{-3}$) |
|---|---|---|---|
| Quartz | 37 | 44 | 2.65 |
| Clay | 21 | 7 | 2.60 |
| Total organic carbon | 2.9 | 2.7 | 1.25 |
| Water | 2.5 | 0 | 1.03 |
| Gas | 0.00013 | 0 | 0.00065 |

When the clay content remained constant, adjustments were made to the TOC content and porosity. The outcomes of the elastic parameters simulated by the rock physics model are presented in Figure 7. It is evident that parameters such as $V_p$, $V_s$, $V_p/V_s$ and P-wave impedance (PI) declined as TOC content increased. This is attributable to the lower bulk and shear moduli of TOC compared to those of sand and mud. The formation of TOC results in a reduction of the $V_p$ and $V_s$ of the reservoir. The $V_p$, $V_s$, $V_p/V_s$, and the PI decreased with increasing porosity. The presence of a low $V_p/V_s$ and low PI may indicate the development of pores.

When porosity was held constant, and clay content and TOC were varied, the resulting elastic parameters predicted by the rock physics model are shown in Figure 8. It is evident that $V_p$, $V_s$, and PI decreased with an increase in clay content, while the $V_p/V_s$ increased with increasing clay content. This indicates that $V_p$ is more sensitive to changes in clay content, and the development of LVM may occur in layers with high $V_p/V_s$ and low PI.

### 3.4. Inversion strategy

#### *3.4.1. Core velocity inversion*

The kerogen macerals in the core were identified, and the percentages of each maceral were quantified to determine the type of organic matter. The detection process was conducted using the fluorescence microscope (Axio Scope A1, ZEISS, Germany). Core samples from the survey area were obtained from qualified drilling. Figure 9 shows the kerogen macerals in the core samples at four depth points. According to the classification and analysis of organic matter types, the type index (TI) of sample A is −53.48, the TI of sample B is −57.35, the TI of sample C is −46.16, and the TI of sample D is −55.96, respectively. The detection results for the four samples showed that the kerogen type was type III.

Total organic carbon, $V_p$, and $V_s$ measurements were also based on authentic samples. Table 3 lists the measurements for nine core samples. Prior to TOC determination, 100 mg of powder was meticulously weighed using an analytical balance. The samples were then immersed in a dilute hydrochloric acid solution (analytically pure) for at least 12 h. Thereafter, they were rinsed with deionized water to achieve neutrality. The samples were desiccated in an oven at 60 °C. The TOC of the samples was determined using a carbon–sulfur analyzer (CS744, LECO, America). The ultrasonic frequency band was used to measure both the longitudinal wave and shear wave velocities. The device under consideration consisted of a pulse generator, a digital oscilloscope, a longitudinal- and transverse-wave probe, a microcomputer, and a processing algorithm. The ultrasonic response of the core was measured using the transmission method, with a transducer frequency of 1,000 kHz. The crossplot of TOC with $V_p$, $V_s$, $R_t$, and $V_p/V_s$ of nine core samples is shown in Figure 10. Testing results show that the $V_p$, $V_s$, $R_t$, and $V_p/V_s$ decreased with increasing TOC, consistent with the rock physics modeling simulation conclusion.

**Table 3. Total organic carbon (TOC), P-wave velocity ($V_p$),**

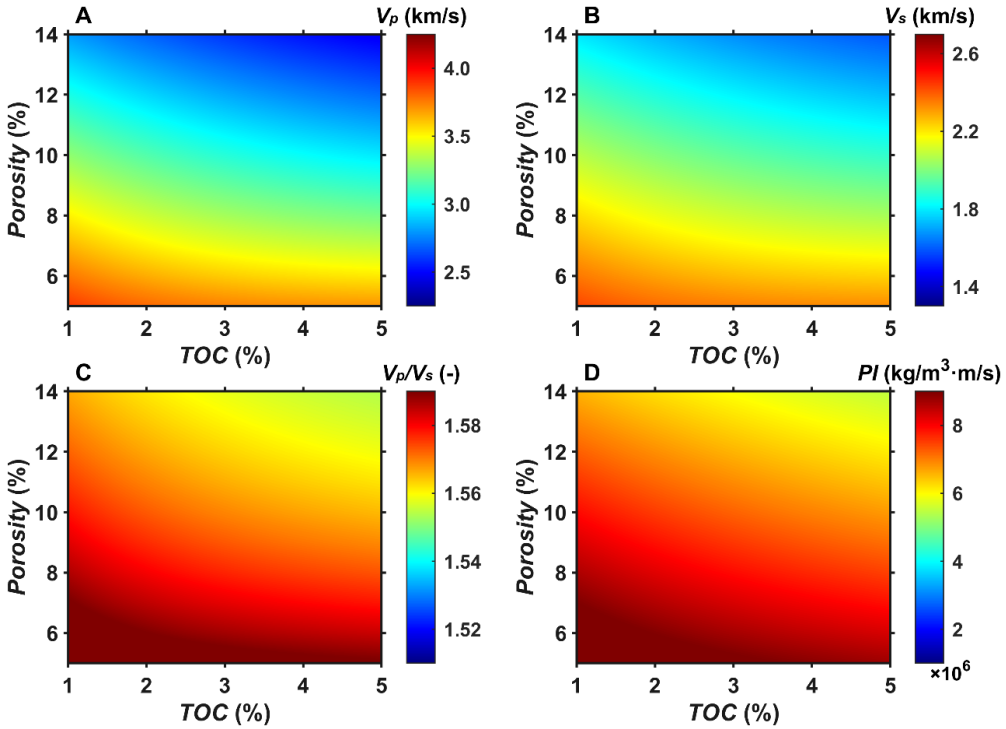



**Figure 7.** The variation of $V_p$, $V_s$, $V_p/V_s$ and PI with TOC and porosity. (A) $V_p$ with different TOC and porosity. (B) $V_s$ with different TOC and porosity. (C) $V_p/V_s$ with different TOC and porosity. (D) PI with a different TOC and porosity.
Abbreviations: PI: P-wave impedance; TOC: Total organic carbon; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

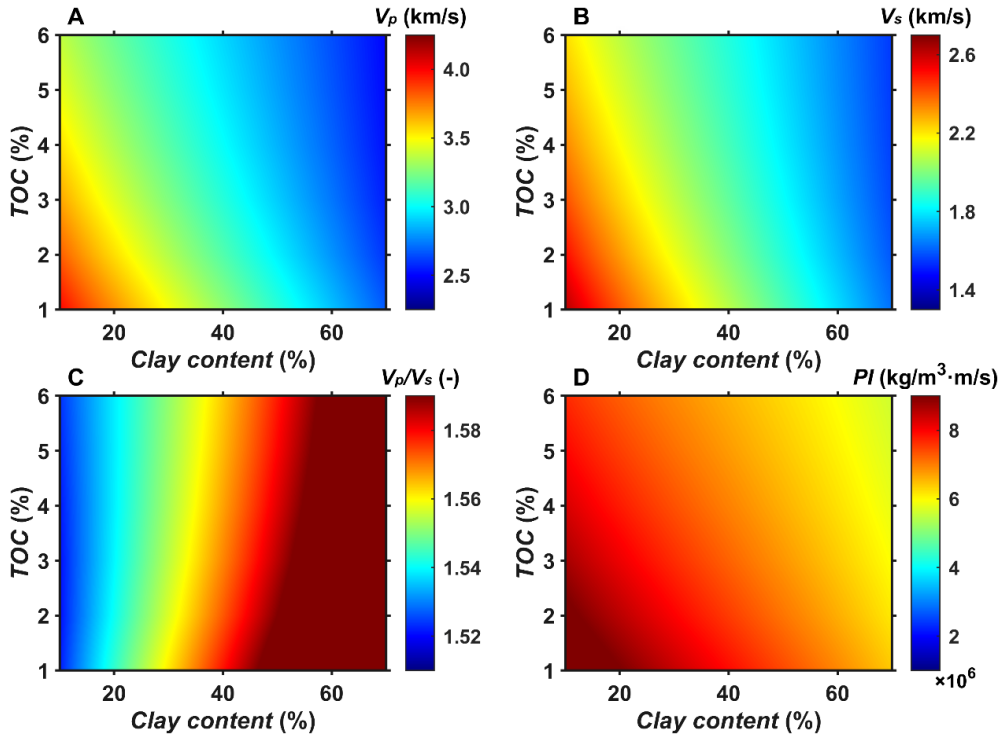



**Figure 8.** The variation of $V_p$, $V_s$, $V_p/V_s$ and PI with clay content and TOC. (A) $V_p$ with different clay content and TOC. (B) $V_s$ with different clay content and TOC. (C) $V_p/V_s$ with different clay content and TOC. (D) PI with different clay content and TOC.
Abbreviations: PI: P-wave impedance; TOC: Total organic carbon; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

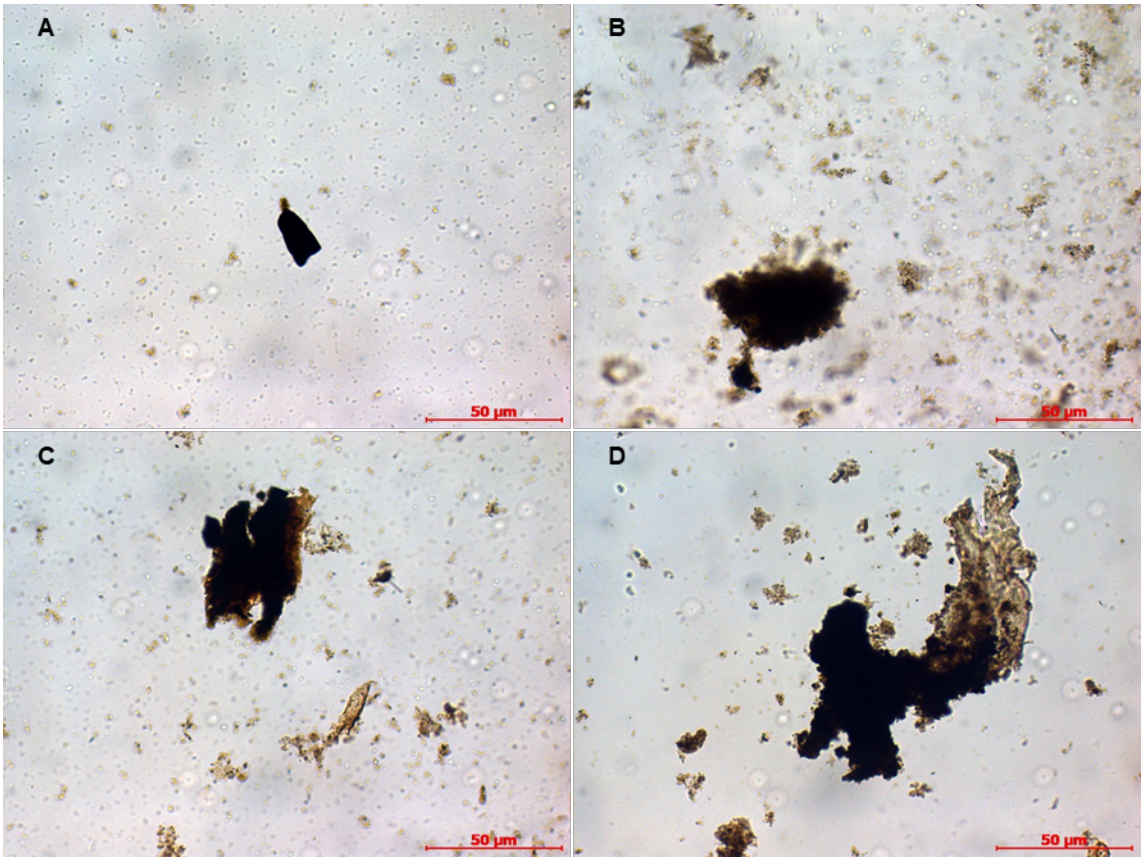



**Figure 9.** Micrograph images of kerogen macerals of four wall core samples from the survey area. (A) Mudstone of the borehole core sample, measured depth at 3,402 m. (B) Mudstone of borehole core sample, measured depth at 3,595 m. (C) Mudstone of borehole core sample, measured depth at 3,769 m. (D) Mudstone of the borehole core sample, measured depth at 3,897 m. Scale bar: 50 μm; magnification: 500×.

**and S-wave velocity ($V_s$) data based on authentic core samples**

| Core sample | TOC (%) | $V_p$ (m/s) | $V_s$ (m/s) |
|---|---|---|---|
| C-1 | 1.17 | 3,119 | 1,696 |
| C-2 | 5.67 | 2,959 | 1,687 |
| C-3 | 2.95 | 3,562 | 1,989 |
| C-4 | 3.50 | 3,249 | 1,770 |
| C-5 | 0.41 | 3,699 | 2,206 |
| C-6 | 0.63 | 3,685 | 1,993 |
| C-7 | 2.09 | 3,471 | 1,785 |
| C-8 | 1.00 | 3,161 | 1,587 |
| C-9 | 1.59 | 2,930 | 1,604 |

The velocities of the nine samples were predicted to assess the validity of the aforementioned rock physics modeling method. As illustrated in Figures 11 and 12, the predicted and actual velocities were compared, with or without TOC. The ordinate denotes the predicted velocity, while the abscissa represents the actual velocity. It is evident that in the absence of kerogen, the predicted $V_p$ and $V_s$ exhibited higher than the actual velocities. Conversely, the predicted results that account for TOC showed greater alignment with the actual values.

### *3.4.2. Logging velocity inversion*

The rationality of the rock physics model developed in this study was verified using actual logging data of Well X in an exploration area. Inversion of $V_s$ can be defined as a nonlinear optimization problem. It is evident that simulated annealing is a reliable nonlinear inversion method.[32–35] In the current study, the actual velocity $\tilde{V}_p$, inverted velocity $V_p$, actual $\tilde{V}_p/\tilde{V}_s$ and inverted $V_p/V_s$ were used to develop the inversion objective function as $g = \arg\min\left(\left\|\tilde{V}_p - V_p\right\|_2 + \left\|\tilde{V}_p/\tilde{V}_s - V_p/V_s\right\|_2\right)$. The $V_s$ was calculated by inversion and compared with the measured velocity to evaluate the rationality of inversion and the effectiveness of the rock physics model. To predict inversions, it is necessary to consider several factors, including logging data and rock physical properties. The logging process involves measuring and interpreting parameters such

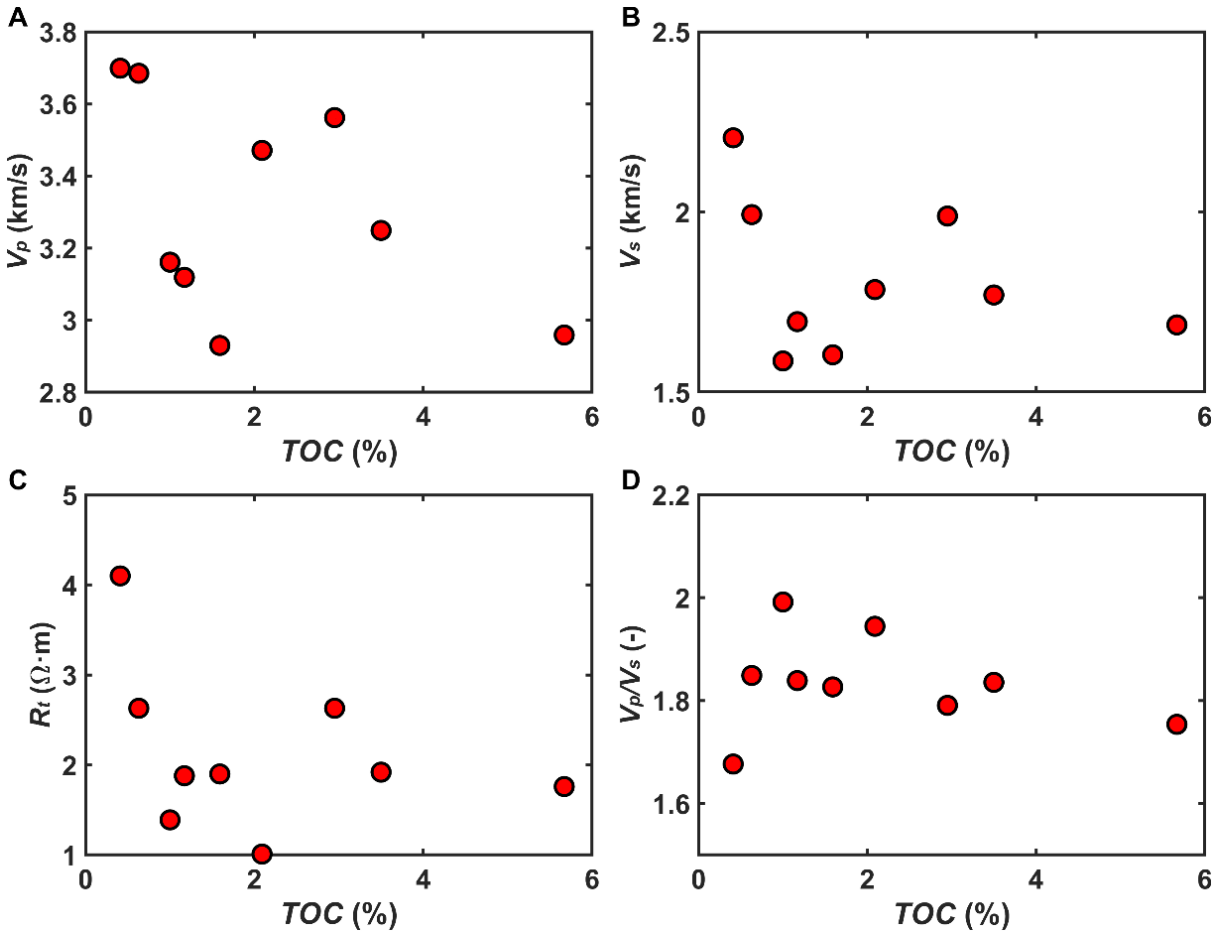



**Figure 10.** The cross-plots of the TOC with $V_p$, $V_s$, $R_t$, and $V_p/V_s$. (A) TOC and $V_p$. (B) TOC and $V_s$. (C) TOC and $R_t$. (D) TOC and $V_p/V_s$.
Abbreviations: $R_t$: Resistivity; TOC: Total organic carbon; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

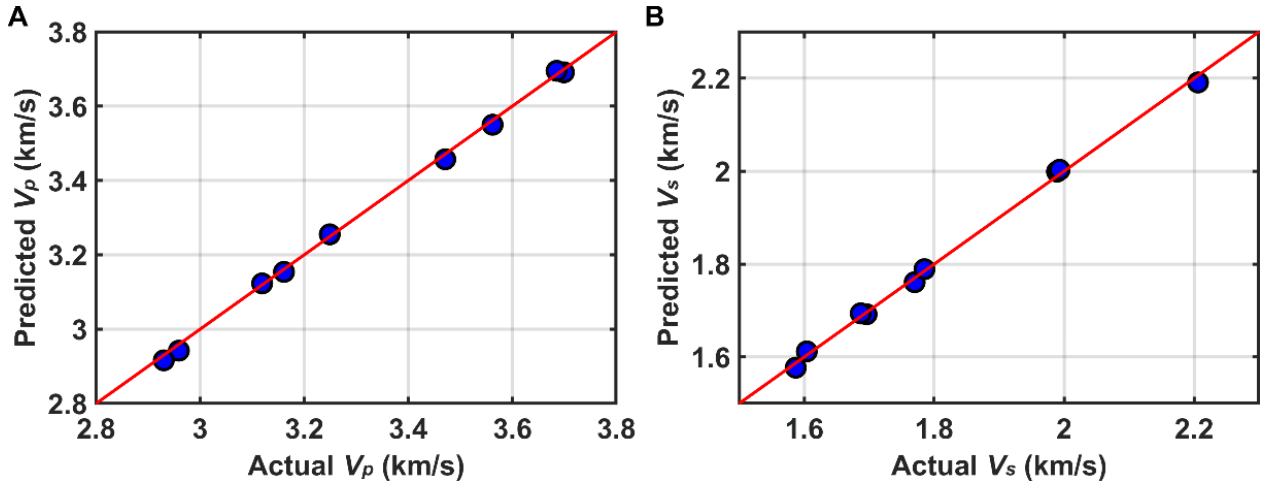



**Figure 11.** The cross-plots of the actual and predicted (A) $V_p$ and (B) $V_s$
Abbreviations: $V_p$: P-wave velocity; $V_s$: S-wave velocity.

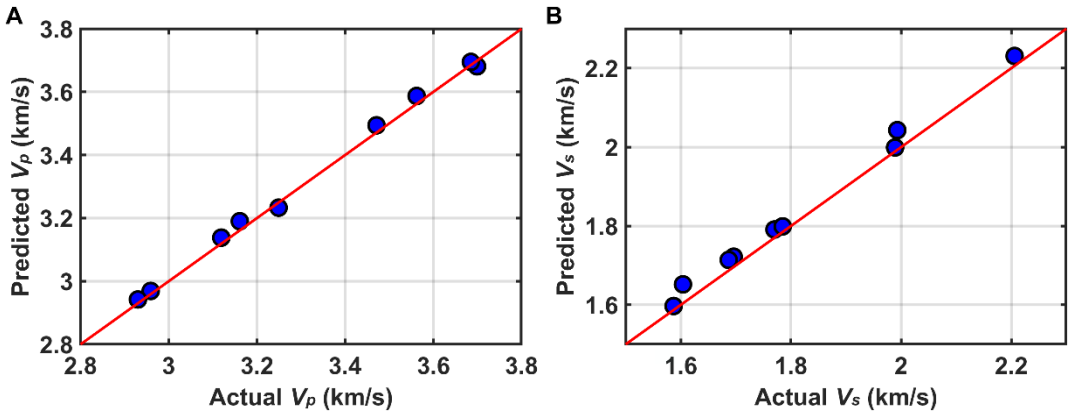



**Figure 12.** The cross-plots of the actual and predicted (A) $V_p$ and (B) $V_s$, without the effect of total organic carbon
Abbreviations: $V_p$: P-wave velocity; $V_s$: S-wave velocity.

as mineral content, fluid saturation, porosity, and rock density.[36,37] These parameters can be obtained through direct measurement and logging interpretation. The logging data indicate that Well X was characterized by a medium-deep, low-porosity, and low-permeability clastic rock reservoir. The rock's mineralogical composition was dominated by quartz and clay, with occasional traces of calcite and mica. The porosity curve was interpreted from density logging data, while the saturation curve was calculated from resistivity logging data using the Archie formula. All interpretation conclusions were calibrated. The rock's physical parameters included the bulk modulus and shear modulus of each component, which were determined by the rock physical test results of the actual coring.

Figure 13 presents the comparison between the predicted and actual $V_p$ and $V_s$ in Well X. The lithology color marks—yellow and black—denote sandstone and mudstone, respectively. In the longitudinal and transverse wave curves, the ordinate denotes the depth, the abscissa signifies the velocity, the blue curve represents the inversion result, and the red curve corresponds to the measured velocity. Figure 14 presents the statistical analysis of inversion velocity error. The inversion $V_p$ error was less than 5%, the $V_s$ error was less than 10%, and the velocity error was small and generally presented a normal distribution also as shown in Figure 15. This indicates the applicability of the constructed rock physics model in the LVM formation in the exploration area.

The preceding analysis demonstrated the efficacy of the rock physics model in describing LVM containing organic matter, and the model-estimated elastic parameters were highly reliable.

## 4. Field examples

In the inversion strategy discussed in this study, the PI and velocity ratio were inverted using pre-stack seismic partial-stack gather data. This work was completed using the Jason software (version 24.2, GeoSoftware, USA). Prior to pre-stack inversion, it is imperative to preserve seismic data amplitude by applying fine-wavefront diffusion compensation, pre-stack denoising, and multiple removal, among other processes. Furthermore, it was assumed that the effects of interlayer multiples and anisotropy subsequent to processing can be disregarded. The wavelets used in the inversion were obtained via well-seismic calibration, and the low-frequency models were interpolated from well data.

For field seismic data from the Yinggehai Basin in the South China Sea, conventional pre-stack simultaneous inversion was previously regarded as a productive technique for differentiating sandstone and mudstone. However, implementing this strategy is challenging during the LVM development interval, which is characterized by the presence of organic matter. Figure 16 presents the seismic sections of the target area. Figure 16A–C corresponds to the 0–10°, 10–20°, and 20–30° (near, middle, and far) stack seismic sections, respectively.

The inversion results of PI and $V_p/V_s$ are shown in Figure 17. Figure 17A,B represents the PI section and the $V_p/V_s$ section, respectively. The red rectangular regions denote the GS, while the black rectangular region indicates the LVM layer.

It is noteworthy that in the inversion profiles of PI and the $V_p/V_s$, although the velocity ratio exhibited weaker lateral continuity compared to the impedance, distinct low-value

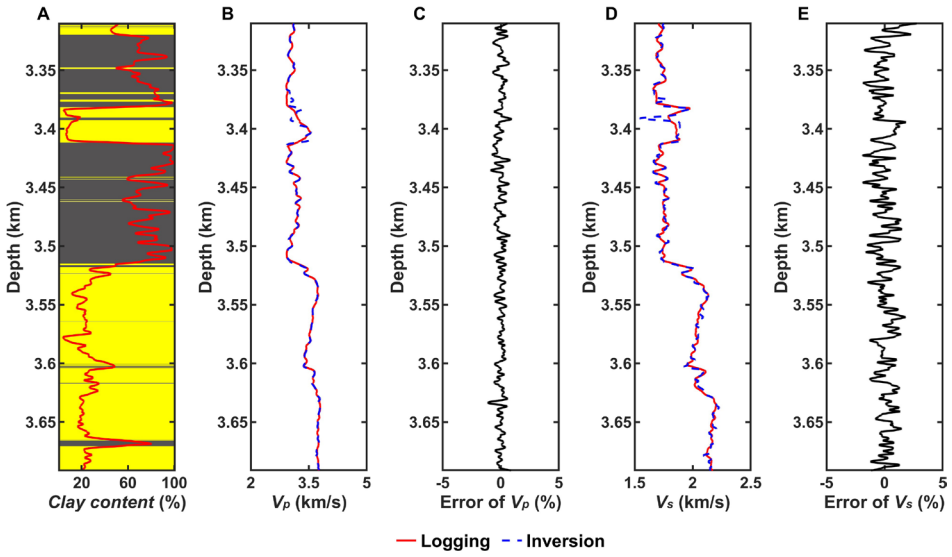



**Figure 13.** The comparison of the predicted and the actual $V_p$ and $V_s$ of Well X. (A) Lithology and clay content. (B) $V_p$ of logging and inversion data. (C) Error of $V_p$. (D) $V_s$ of logging and inversion data. (E) Error of $V_s$.
Abbreviations: $V_p$: P-wave velocity; $V_s$: S-wave velocity.

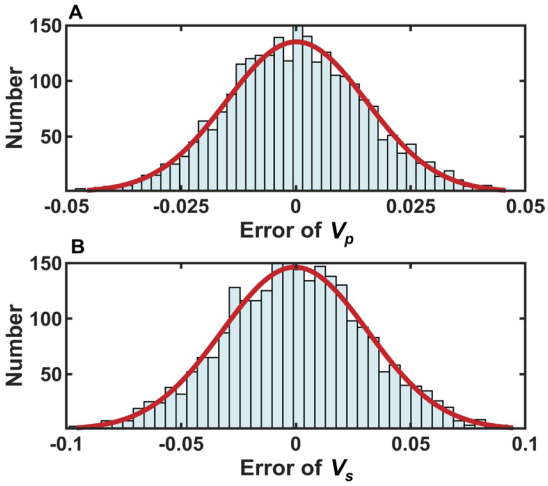



**Figure 14.** Error distribution between the measured and predicted (A) $V_p$ and (B) $V_s$
Abbreviations: $V_p$: P-wave velocity; $V_s$: S-wave velocity.

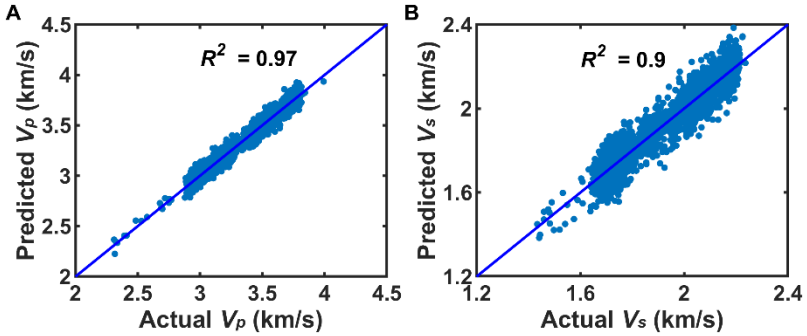



**Figure 15.** Cross-plots of the measured and predicted P and S wave velocities. (A) the $V_p$; and (B) the $V_s$.
Abbreviations: $V_p$: P-wave velocity; $V_s$: S-wave velocity.

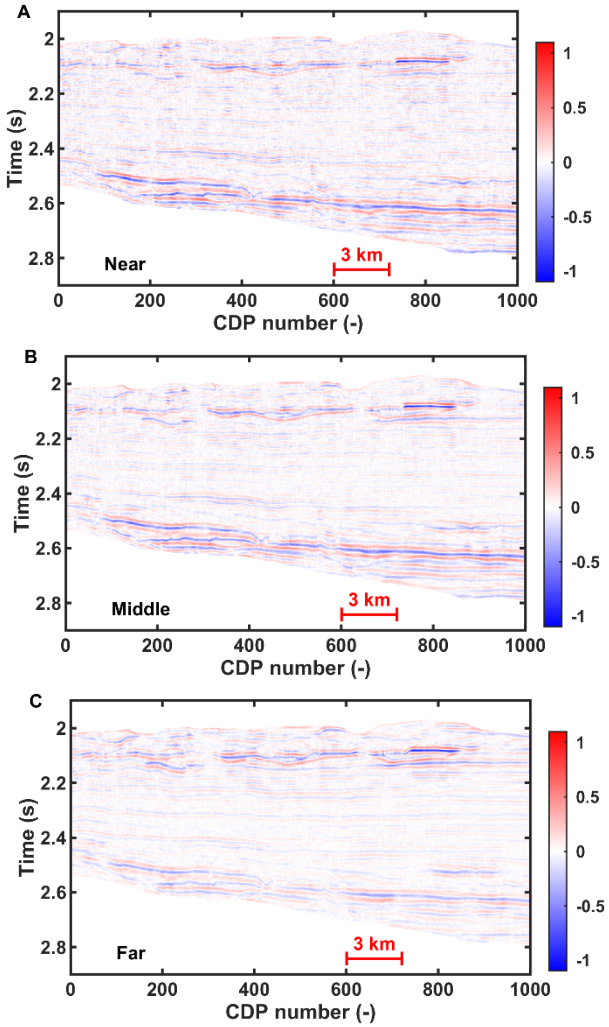



**Figure 16.** Partial angle stacking seismic profiles. (A) Near, stacking angles of 0–10°. (B) Middle, stacking angles of 10–20°. (C) Far, stacking angles of 20–30°.
Abbreviation: CDP: Common depth point.

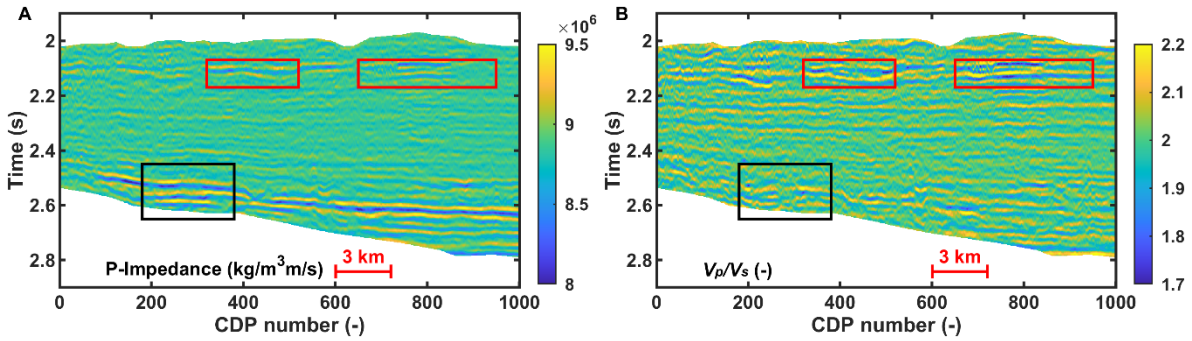



**Figure 17.** Pre-stack inversion results of elastic parameters. (A) P-impedance. (B) $V_p/V_s$.
Abbreviations: CDP: Common depth point; $V_p$: P-wave velocity; $V_s$: S-wave velocity.

anomalies can still be observed within the interval marked by the black rectangle. The consensus within the geophysical community is that impedance serves as an effective indicator for distinguishing sandstone from mudstone, while $V_p$ demonstrates higher sensitivity to variations in gas saturation. Consequently, Figure 17 illustrates a key finding of this study that PI may misinterpret LVM as GS, and the low values present in the $V_p/V_s$ profile further exacerbate this risk of misinterpretation. It is evident that all samples exhibited the characteristics of low PI and low $V_p/V_s$, which carries the potential risk of misinterpreting the LVM formation as GS.

In light of the aforementioned analysis and the novel rock physics model, the simulated annealing nonlinear inversion strategy was employed to facilitate the seismic inversion of petrophysical parameters for LVM reservoirs.[38–43] A well-connected line was selected, and the elastic parameters of the line were obtained by pre-stack inversion. These were then used as input data for physical parameter inversion. The initial inversion value was also generated by well interpolation, as shown in Figure 18. Inverted porosity, clay content, and water saturation are demonstrated in Figure 19. As shown in the seismic profile, there are strong-amplitude regions at travel times of 2.6 s and 2.1 s. It is challenging to circumvent LVM interference using conventional AVO attributes and pre-stack simultaneous inversion. As illustrated in Figure 19, the red rectangular regions denote the GS formation, while the black rectangular region indicates the LVM formation. The inversion results indicate that both porosity and gas saturation were elevated near 2.1 s, while there was an absence of discernible gas-bearing anomalies in the LVM section near 2.6 s. The inversion results provide a superior reflection of the actual drilling understanding, circumvent the LVM risk associated with conventional elastic parameters when predicting gas-bearing sweet spots, and facilitate the decoupling of GS and LVM.

The reliability of the inversion process can be evaluated by extracting the beside-borehole trace data from the inversion results and comparing them with the actual well data. Figure 20 compares the inversion results with the true values of porosity, clay content, and gas saturation. The reliability of the prediction results obtained using the rock physics model and inversion strategy in our study is well supported. The findings of this study demonstrate that, in the absence of consideration for the evident anisotropy, the petrophysical parameters of the formation can be reasonably inverted by employing the rock physics modeling and seismic inversion method delineated in this study. This is because the mudstone in the study area lacks a discernible directional arrangement of cracks or laminae, thereby allowing the influence of anisotropy on velocity to be reasonably ignored.[44]

## 5. Discussion

It is considered reliable for identifying sandstone and mudstone using rock physics intersection analysis and pre-stack simultaneous inversion. Moreover, it can be used to identify GS and dry sandstone. However, the petrophysical superposition of mechanism-containing LVM and GS in the Yinggehai Basin is significant, and the interpretation scheme for elastic parameters increases the risk of misjudgment of LVM as GS.

The organic matter content of mudstone significantly influences its elastic properties, thereby reducing its velocity. This effect is accompanied by a strong LVM amplitude, a phenomenon that cannot be disregarded in rock physics modeling. The establishment of a rock physics model for the LVM reservoir in the Yinggehai Basin is predicated on analysis of relevant geological data and seismic rock physics characteristics. The model was developed with consideration for the influence of TOC. The influence of TOC on seismic elastic properties was considered, and the contributions of TOC, porosity, and clay content to velocities were analyzed. The simulation demonstrated that the $V_p$, $V_p/V_s$, and PI decreased with an increase in TOC, thereby elucidating the microscopic rock-physics mechanism of mudstone low impedance anomaly.

Notably, although the content of organic matter in mudstone was higher, the overall content was lower than that of shale source rocks. In contrast to the substantial vertical transverse isotropy exhibited by shale, the organic-rich LVM predominantly manifested as a massive structure with minimal brittle mineral content. The fracturing process is challenging, and the current geological survey does not consider its horizon to be conducive to oil and gas reservoir formation. However, in this study, the viscosity of organic matter significantly affected the propagation velocity of seismic waves.

It is important to note that kerogen's viscosity can cause velocity dispersion and wave energy attenuation, leading to differences across acoustic, logging, and seismic frequency bands. The extant research demonstrates that the velocity of seismic waves in the low-frequency range is typically marginally lower than that of the logging and ultrasonic frequency bands. The associated modeling can account for wave-induced fluid flow mechanisms or viscoelastic mechanisms. However, when taken in conjunction with the Biot flow theory and squirt flow theory, the Biot attenuation peak of fast longitudinal waves will shift to the high frequency band with an increase in fluid viscosity, while the squirt flow attenuation peak will shift to the

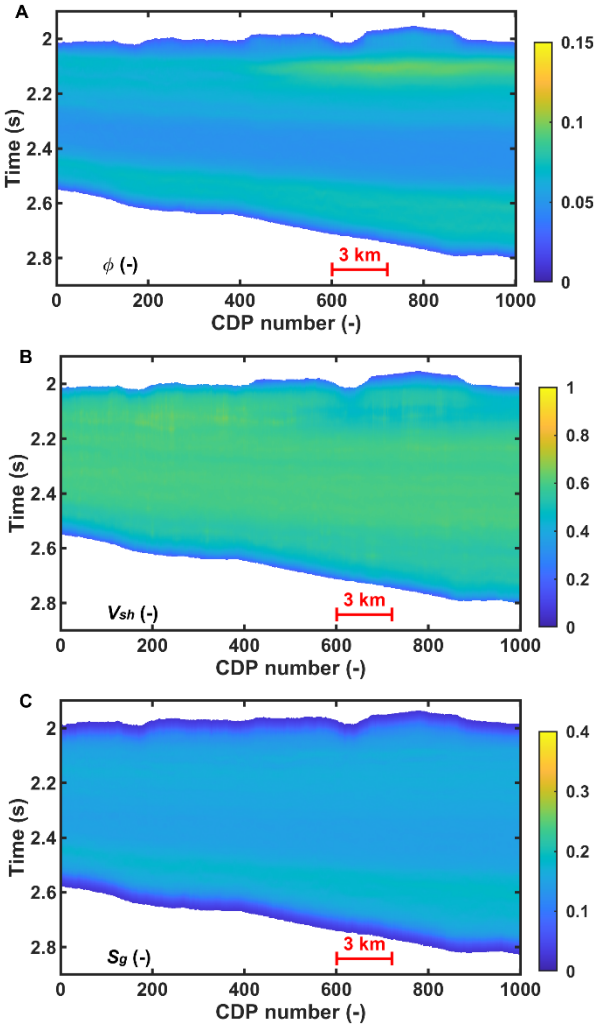



**Figure 18.** Initial model profile of (A) porosity ( $\phi$ ), (B) clay content ( $V_{sh}$ ), and (C) gas saturation ( $S_g$ )
Abbreviation: CDP: Common depth point.

low frequency band with an increase in fluid viscosity. The incorporation of this attenuation mechanism into rock physics modeling introduces additional parameters, such as tortuosity and crack density. The accurate and precise articulation of these parameters is challenging. It is therefore evident that the modeling scheme described in this study is based on the equivalent medium theory, and the attenuation and dispersion of kerogen pairs can be discussed using multi-band rock physics experimental data.

As the elastic modulus of kerogen is considerably lower than that of background minerals such as quartz, in this study, kerogen and background minerals were not treated as equivalent to the background matrix. Rather, they were part of the pore filler, and their low-frequency characteristics were maintained. Organic matter may be found within micro-cracks, intergranular pores, or biological cavity pores. However, the development pattern of kerogen in rock media is more complex. Adsorption of organic matter on the surface and within the pores of mineral particles (particularly clay minerals) occurs in an amorphous form, constituting the primary manifestation of the dispersion pattern at the micro level. Under a microscope, kerogen and mineral particles were observed to be intermingled, forming an organic–mineral complex. This complicates the process of distinguishing individual kerogen particles, and the rock physics modeling method for this distribution pattern warrants further discussion in future work.

## 6. Conclusion

The seismic elastic characteristics of organic-rich LVM have

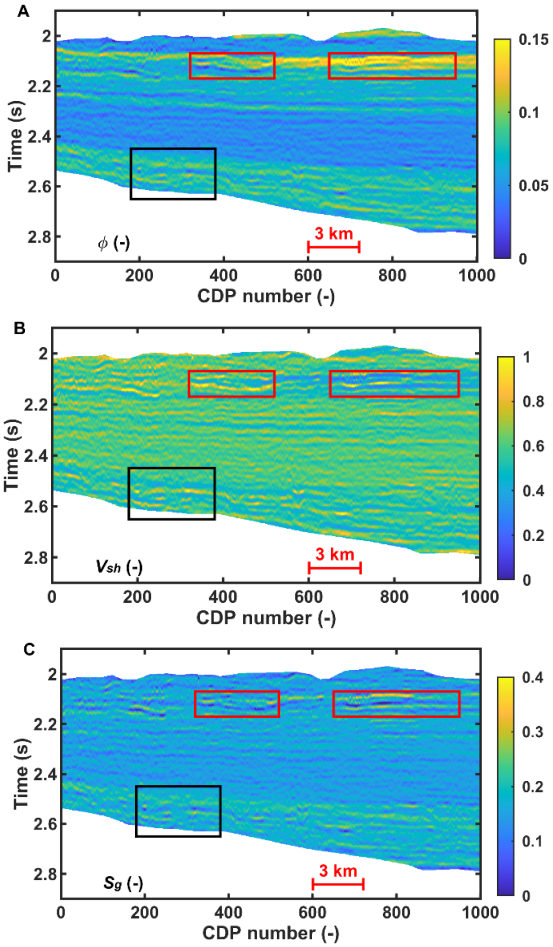



**Figure 19.** Profile of (A) porosity ( $\phi$ ), (B) clay content ( $V_{sh}$ ) and (C) gas saturation ( $S_g$ ) from inversion
Abbreviation: CDP: Common depth point.

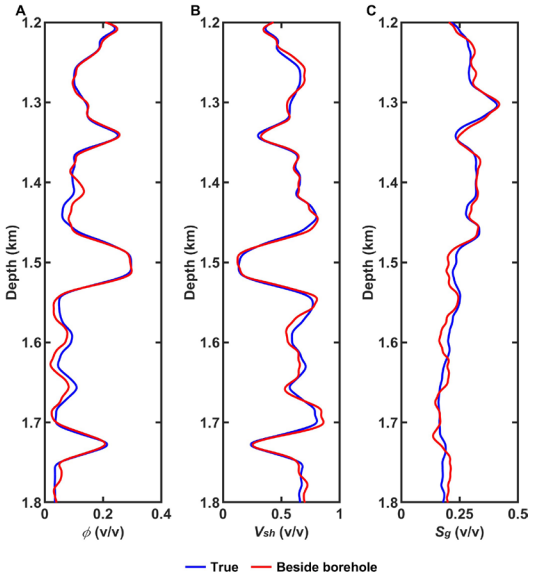



**Figure 20.** Comparison of the true data and the predicted borehole results of porosity $\phi$, clay content $V_{sh}$ and gas saturation $S_g$, for which the true data are in blue, and the predicted results are in red. (A) Porosity $\phi$. (B) clay content $V_{sh}$. (C) gas saturation $S_g$.

been shown to be analogous to those of GSs. The presence of kerogen in the mudstone of the Yinggehai Basin has been demonstrated to be the cause of the pronounced amplitude observed. The present study used the rock physics modeling scheme of the SCA model and the Ciz–Gassmann equation to predict the $V_p$ and $V_s$ of LVM containing organic matter. This theoretical foundation facilitates the identification of LVM through seismic inversion. The methodology employed was founded on the principles of rock physics, specifically the $V_p$ and $V_s$ equations. This approach enables the robust inversion of porosity, clay content, and saturation. The subsequent data processing serves to validate the method's viability for practical applications and, concurrently, provides a seismic inversion method for identifying GS and LVM from seismic data. The method was applied to the exploration target, enabling quantitative seismic predictions of porosity, clay content, and saturation using a rock physics model. The inversion results were in good agreement with the drilling results, which verifies the effectiveness of the method. In the context of reservoir seismic identification, it is imperative to enhance the precision of inversion predictions by leveraging seismic and logging data alongside the evaluation of formation lithology. Consequently, it is imperative to establish a direct relationship between the seismic reflection coefficient and the physical parameters in subsequent work. Moreover, developing a rational pre-stack direct inversion method is essential to eliminate cumulative error in indirect inversion. This will facilitate a more effective execution of reservoir seismic identification and evaluation.

## Acknowledgments

The authors express their gratitude to Dr. Kun Lang, Dr. Songhe Yu, and Professor Shiyou Liu for their discussions.

## Funding

This study was financially supported by the National Oil and Gas Major Project of China, "Amplitude preserving imaging technology with high precision and multiple geophysical field exploration technology for deep-water complex geological conditions in South China Sea" (Grant No. 2025ZD1402706).

## Conflict of interest

The authors declare that they do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

## Author contributions

*Conceptualization:* Jinpeng Li, Yanhua Xie
*Formal analysis:* Yanhua Xie, Jinpeng Li, Wanyuan Sun
*Investigation:* Yanhua Xie, Wanyuan Sun, Yayun Dang
*Methodology:* Jinpeng Li, Yanhua Xie, Fang Li
*Visualization:* Yanhua Xie, Wanyuan Sun
*Writing–original draft:* Yanhua Xie, Wanyuan Sun
*Writing–review & editing:* Yanhua Xie, Jinpeng Li, Yayun Dang

## Availability of data

The data in this paper is subject to the following licenses/ restrictions: National confidential data. Requests to access these datasets should be directed to the corresponding author.

## References

1. Guo XX, Hu L, Wu YY, Xiong XF, Yu HW, Xia L. Gas generation potential of Oligocene source rocks in Yinggehai Basin. *Nat Gas Ind B*. 2024;11(6):728-738.

   doi: 10.1016/j.ngib.2024.11.006

2. Pei JX, Yu JF, Wang LF, Hao DF, Liu F. Key challenges and strategies for the success of natural gas exploration in mid-deep strata of the Yinggehai Basin. *Acta Pet Sin*. 2011;32(4):573-579. [in Chinese].

   doi: 10.7623/syxb201104003

3. Wu KQ, Pei JX, Hu L, Hu GW, Tang SM, Zhao H. Accumulation model and exploration direction of medium-large gas fields in Yinggehai Basin. *Acta Pet Sin*. 2023;44(12):2200-2216. [in Chinese].

   doi: 10.7623/syxb202312012

4. Zhang XY, Fan CW, Alves TM, *et al*. Hydrocarbon generation in source rocks with low organic matter as the primary mechanism for overpressure development: A case study from the Yinggehai Basin, South China Sea. *J Geol Soc*. 2025;183(1):jgs2024-157.

   doi: 10.1144/jgs2024-157

5. Xie YH, Li XS, Tong CX, Liu P, Wu HZ, Huang ZL. High temperature and high pressure gas enrichment condition, distribution law and accumulation model in central diaper zone of Yinggehai basin. *China Offshore Oil Gas*. 2015;27(4):1-12. [in Chinese].

   doi: 10.11935/j.issn.1673-1506.2015.04.001

6. Yang GX, Yin HW, Jia D, Wang HB, Wang W, Xu WQ. New insights into the structure of the Yinggehai Basin and its tectonic implications, South China Sea: Evidence from scaled physical models. *Basin Res*. 2024;36(4):e12888.

   doi: 10.1111/bre.12888

7. Xie YH, Gao YD. Recent domestic exploration progress and direction of CNOOC. *China Pet Explor*. 2020;25(1):20-30.

[in Chinese].

doi: 10.3969/j.issn.1672-7703.2020.01.003

8. Xie YH. China National Offshore Oil Corporation: Major achievements in oil and gas exploration of CNOOC in the 13th Five-Year Plan period and prospects in the 14th Five-Year Plan period. *China Pet Explor.* 2021;26(1):43-54. [in Chinese].

doi: 10.3969/j.issn.1672-7703.2021.01.004

9. Tong CX, Wang ZF, Li XS. Pooling conditions of gas reservoirs in the Dongfang 1-1 Gas Field, Yinggehai Basin. *Nat Gas Ind.* 2012;32(8):11-15. [in Chinese].

doi: 10.3787/j.issn.1000-0976.2012.08.003

10. Li F, Deng Y, Hu L, Zhou F. Porosity prediction by geophysics technology at low-porosity and low-permeability reservoir of the Yinggehai Basin. *Bull Geol Sci Technol.* 2022;41(4):84-90. [in Chinese].

doi: 10.19509/j.cnki.dzkq.2021.0060

11. Grana D, Rossa ED. Della Rossa E. Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics*. 2010;75(3):O21-O37..

doi: 10.1190/1.3386676

12. Russell BH, Gray D, Hampson DP. Linearized AVO and poroelasticity. *Geophysics*. 2011; 76(3):C19-C29.

doi: 10.1190/1.3555082

13. Yin XY, Cui W, Zong ZY, Liu XJ. Petrophysical property inversion of reservoirs based on elastic impedance. *Chin J Geophys.* 2014;57(12):4132-4140. [in Chinese].

doi: 10.6038/cjg20141224

14. Zong ZY, Yin XY, Wu GC. Fluid identification method based on compressional and shear modulus direct inversion. *Chin J Geophys*. 2012;55(1):284-292. [in Chinese].

doi: 10.6038/j.issn.0001-5733.2012.01.028

15. Luo C, Ba J, Guo Q. Probabilistic seismic petrophysical inversion with statistical double-porosity Biot-Rayleigh model. *Geophysics*. 2023; 88(3):M157-M171.

doi: 10.1190/geo2022-0288.1

16. Li K, Yin X, Zong Z. Facies-constrained prestack seismic probabilistic inversion driven by rock physics. *Sci China Earth Sci*. 2020;63(6):822-840.

doi: 10.1007/s11430-019-9578-1

17. Ba J, Ai ZJ, Carcione JM, Pang MQ, Yan XF, Chen X. Estimation of pore structure for heterogeneous reservoirs based on the theory of differential poroelasticity. *Pure Appl Geophys.* 2024, 181(7):2131-2147.

doi: 10.1007/s00024-024-03510-3

18. Li H, Zhang J, Pan H, Gao Q. Nonlinear simultaneous inversion of pore structure and physical parameters based on elastic impedance. *Sci China Earth Sci*. 2021;64(6):977-991.

doi: 10.1007/s11430-020-9773-8

19. Carcione JM. A model for seismic velocity and attenuation in petroleum source rocks. *Geophysics*. 2000;65(4):1080-1092.

doi: 10.1190/1.1444801

20. Zhang GZ, Chen JJ, Chen HZ, Ma ZG, Li CC, Yin XY. Prediction for in-situ formation stress of shale based on rock physics equivalent model. *Chin J Geophys*. 2015;58(6):2112-2122. [in Chinese].

doi: 10.6038/cjg20150625

21. Ba J, Xu W, Fu LY, Carcione JM, Zhang L. Rock anelasticity due to patchy-saturation and fabric heterogeneity: A double-double porosity model of wave propagation. *J Geophys Res Solid Earth.* 2017;122(3):1949-1976..

doi: 10.1002/2016JB013882

22. Dutilleul J, Bourlange S, Géraud Y, Stemmelen D. Porosity, pore structure, and fluid distribution in the sediments entering the northern Hikurangi margin, New Zealand. *J Geophys Res Solid Earth*. 2020;125(11):1-21. .

doi: 10.1029/2020JB020330

23. Hill R. The elastic behavior of crystalline aggregate. *Proc Phys Soc.* 1952;65(5):349-354.

doi: 10.1088/0370-1298/65/5/307

24. Berryman JG. Long-wavelength propagation in composite elastic media I. Spherical inclusions. *J Acoust Soc Am.* 1980;68(6):1809-1819.

doi: 10.1121/1.385171

25. Berryman JG. Long-wavelength propagation in composite elastic media II. Ellipsoidal inclusions. *J Acoust Soc Am*. 1980;68(6):1820-1831.

doi: 10.1121/1.385172

26. Ciz R, Shapiro SA. Generalization of Gassmann's equations for porous media saturated with a solid material. *Geophysics*. 2007;72(6):A75-A79.

doi: 10.1190/1.2772400

27. Brie A, Pampuri F, Marsala AF, Meazza O. Shear sonic interpretation in gas-bearing sands. In: *SPE Annual Technical Conference and Exhibition*. SPE Annual Technical Conference and Exhibition; October 22–25, 1995; Dallas, Texas. Society of Petroleum Engineers (SPE); 1995.

doi: 10.2118/30595-ms

28. Gassmann F. Elastic waves through a packing of spheres. *Geophysics*. 1951;16(4):673-685.

doi: 10.1190/1.1437718

29. Yu SH, Zong ZY, Li K. Rock physics and seismic reflectivity parameterization for amplitude-variation-with-angle inversion of low-maturity source rocks. *Geophysics*. 2023;88(6):R703-R719.

doi: 10.1190/geo2022-0734.1

30. Zong ZY, Yu SH, Tan H, Yin XY, Li K. Nonlinear seismic inversion of total organic carbon for hydrocarbon source rocks. *Geophysics*. 2025;90(5):MR375-MR390.

doi: 10.1190/geo2024-0288.1

31. Zhao LX, Qin X, Han DH, Geng JH, Yang ZF, Cao H. Rock-physics modeling for the elastic properties of organic shale at different maturity stages. *Geophysics*. 2016;81(5):D527-D541.

doi: 10.1190/geo2015-0713.1

32. Zhang JJ, Li HB, Yao FC. Rock critical porosity inversion and S-wave velocity prediction, *Appl Geophys*. 2012;9(1):57-64.

doi: 10.1007/s11770-012-0314-8

33. Yang YH, Yin XY, Gao G, Gui ZX, Zhao B. Shear-wave velocity estimation for calciferous sandy shale formation. *J Geophys Eng.* 2019;16(1):105-115.

doi: 10.1093/jge/gxy009

34. Wang JL, Wu SG, Zhao LX, Wang WW, Wei JG, Sun J. An effective method for shear-wave velocity prediction in sandstones. *Mar Geophys Res*. 2019;40(4):655-664.

doi: 10.1007/s11001-019-09396-4

35. Zhang L, Ba J, Carcione JM, Wu CF. Seismic wave propagation in partially saturated rocks with a fractal distribution of fluid-patch size. *J Geophys Res Solid Earth*. 2022;127(2):e2021JB023809.

doi: 10.1029/2021JB023809

36. Pang MQ, Ba J, Carcione JM, Picotti S, Zhou J, Jiang R. Estimation of porosity and fluid saturation in carbonates from rock-physics templates based on seismic Q. *Geophysics*. 2019;84(6):M25–M36.

doi: 10.1190/geo2019-0031.1

37. Yin XY, Zong ZY, Wu GC. 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

38. Chapman S, Tisato N, Quintal B, Holliger K. Seismic attenuation in partially saturated Berea sandstone submitted to a range of confining pressures. *J Geophys Res Solid Earth*. 2016;121(3):1664-1676.

doi: 10.1002/2015JB012575

39. Bosch M, Mukerji T, Gonzalez EF. Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A Review. *Geophysics*. 2010;75(5):A165-A176.

doi: 10.1190/1.3478209

40. Zhu DH, Gibson R. Seismic inversion and uncertainty quantification using transdimensional Markov Chain Monte Carlo method. *Geophysics*. 2018;83(4):R321-R334.

doi: 10.1190/geo2016-0594.1

41. Kolbjørnsen O, Arild B, Ragnar H, Per R, Abel Onana N, Eyvind A. Bayesian seismic inversion for stratigraphic horizon, lithology, and fluid prediction. *Geophysics*. 2020;85(3):R207-21.

doi: 10.1190/geo2019-0170.1

42. Wang C, Yin C, Shi XW, *et al*. Direct prediction method of fracturing ability in shale formations based on pre-stack seismic inversion. *J Seism Explor.* 2022;31(5):407-424.

43. Zhou XY, Gan LD, Yang TQ, Jiang XY, Wei LL. S-wave velocity estimation and reservoir type identification based on partially connected porosity model of carbonate reservoir. *J Seism Explor*. 2023;32(1):51-65.

44. Sharifi J, Amiri M, Fakhar M, Hafezi Moghaddas N. An experimental study on variation of pore pressure and wave velocity in carbonate rocks: impact of depositional facies and diagenesis. *Geomech Geophys Geo-Energy Geo-Resour.* 2026;12(8):1-31.

doi: 10.1007/s40948-025-01062-x

ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Seafloor reflector imaging in 2D seismic data through muting of out-of-plane signals in the Ulleung Basin, East Sea

**Ganghoon Lee[1], Changyoon Lee[1], Junseok Kwon[1], and Snons Cheong[2]***

[1]Resource Exploration & Development Research Division, Korea Institute of Geoscience and Mineral Resources, Daejeon, Yuseong-gu, Republic of Korea

[2]Geo-Environment Research Division, Korea Institute of Geoscience and Mineral Resources, Daejeon, Yuseong-gu, Republic of Korea

## Abstract

Irregular topography can generate out-of-plane signals (OPS) on seismic sections, interfering with the imaging of the true seafloor directly beneath the survey line. While acquiring three-dimensional data or using specialized sensors can mitigate this, these options are often costly or unavailable, especially for legacy surveys. To efficiently remove OPS from two-dimensional (2D) data, this study investigates the validity of using a neural network (NN) for picking and muting. First, we demonstrate the limitation of conventional frequency–wavenumber domain directional filtering due to the kinematic similarity between OPS and true seafloor reflections. Then, we present a workflow that employs a cascade–correlation learning algorithm to identify and mute OPS arrivals before the first break. Unlike data-intensive deep learning techniques that require large training datasets, this lightweight NN is trained on user-picked examples of true seafloor reflections, enabling it to distinguish OPS events arriving from outside the vertical survey plane. Application of this technique to a 2D line acquired near irregular seafloor topography in the Ulleung Basin demonstrates the true seafloor reflector and the removal of false offline signals. Qualitative and quantitative validation against an independent external bathymetric reference both showed a reduction in travel time error compared to the raw data, confirming the effectiveness of the picking results. The results highlight that a cascade–correlation NN-based picking and muting can efficiently suppress OPS in cases of irregular topography on 2D seismic data.

***Keywords:*** Out-of-plane signal; Neural network picking and muting; True seafloor; Ulleung Basin

***Corresponding author:**
Snons Cheong
(snons@kigam.re.kr)

**Citation:** Lee G, Lee C, Kwon J, Cheong S. Seafloor reflector imaging in 2D seismic data through muting of out-of-plane signals in the Ulleung Basin, East Sea. *J Seismic Explor*. 2026;35(2):025470116. doi: 10.36922/JSE025470116

**Received:** November 21, 2025

**Revised:** February 4, 2026

**Accepted:** February 27, 2026

**Published online:** April 30, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Marine seismic surveys are common tools for exploring the Earth's subsurface beneath oceans and seas, with two-dimensional (2D) profiles, in particular, providing valuable insights into geological structures and stratigraphy.[1] These surveys typically involve a vessel towing an energy source and an array of receivers (streamer), allowing efficient acquisition that yields extensive coverage and high-fold data compared to onshore surveys. Despite these operational advantages, seismic data acquired in the marine

setting are often affected by various types of noise from the environment and the acquisition geometry itself, which can interfere with the interpretation of subsurface targets. Standard seismic data processing aims to enhance the desired signals and attenuate noise, producing a clearer image of the subsurface.[2] However, one type of coherent noise, often encountered when surveying areas characterized by irregular seafloor morphology, poses a particular challenge. This challenge stems from energy reflecting off features that are lateral to the intended 2D survey line. These signals, known as out-of-plane signals (OPS),[3] are sometimes also referred to as sideswipe noise.

Out-of-plane signals originate from irregular seafloor topography or subsurface structures outside the vertical plane of the seismic profile but arrive at the receivers, contaminating the seismic section.[4] As this out-of-plane energy is imaged within the 2D section, it generates false reflectors at apparent depths that may be either deeper or shallower than the true reflection point. A single continuous reflector can then appear as a challenging pattern of discontinuous reflections on the final image.[3,5] These shifted events appear as coherent signals on seismic records but do not correspond to true subsurface structures directly under the line, often leading to misinterpretation if not recognized.[6] As Tucker[7] noted, such out-of-plane energy will almost always be present when 2D lines are shot at orientations other than perpendicular to dip or structural trends. In crooked-line surveys, this problem is particularly severe. The scattered midpoint geometry, combined with cross-profile-dipping reflectors, introduces significant travel time distortions known as cross-dip moveout (CDMO), which degrades the quality of standard common midpoint stacking if not properly addressed.[8] From a geometrical perspective, crooked-line surveys can be regarded as a limited swath three-dimensional (3D) investigation, and in principle, 3D prestack migration provides an appropriate means to image reflectivity in the surrounding volume.[9] However, applying 3D processing to crooked-line data is rarely practical, and interpreters are usually constrained to legacy 2D workflows. Accurate imaging of the true seafloor is important, especially when utilizing legacy 2D data to plan new infrastructure such as submarine cables or structure construction, or to site ocean-bottom instruments. In these engineering or site-survey applications, removing OPS becomes crucial for reliable interpretation. Despite its importance, the specific problem of OPS removal has received relatively less attention.

The most direct solution is acquiring full 3D seismic data, which resolves the true 3D geometry and largely eliminates ghost or sideswipe reflections from offline features, except perhaps near the survey edges.[10-13] Consequently, data-processing approaches are typically employed, but these approaches also suffer from significant limitations. However, 3D surveys are more expensive and may not be feasible for regional studies or when working with existing legacy 2D datasets. In 2D data processing, interpreters sometimes resort to manually muting suspected OPS events or attempt to design survey layouts that avoid proximity to strong lateral reflectors. Other processing approaches include specialized techniques such as frequency–wavenumber (F–K) directional filtering or polarization filtering (when multi-component data are available) to attenuate energy that does not propagate in the inline direction.[14] For multi-streamer or swath-acquisition geometries, this out-of-plane energy can sometimes be identified by its characteristic arrival-time differences across the receivers, often producing a distinctive "sawtooth pattern" when data are sorted in the crossline direction. This property enables targeted suppression through array processing techniques such as null steering.[15] More advanced approaches employ iterative processing workflows that correct for CDMO and then recalculate velocities and dip moveout (DMO), though these can produce artifacts such as duplicated reflection events if not applied carefully.[16] To mitigate this, generalized CDMO (GCDMO) methods have been introduced, designed to operate on curved acquisition geometries, thereby reducing such artifacts.[17] 3D effects present such a challenge that even advanced operators like azimuth moveout inherently contain out-of-plane, skewed-saddle characteristics.[18] To address this issue, velocity-independent, slope-based time-domain DMO techniques were developed that use local event slopes to circumvent the iterative cycle between velocity analysis and DMO correction.[19] However, these approaches can be challenging to apply in practice to field survey data. OPS events often exhibit moveouts and wavelet characteristics similar to those of genuine reflections from beneath the survey line, making them difficult to separate using standard filtering techniques. Simple filtering also carries the risk of removing valid dipping events or diffractions that are part of the true subsurface image.

Distinguishing OPS from the true seafloor reflection based solely on visual inspection or conventional filtering is often difficult. These challenges highlight the need for a more targeted approach capable of accurately identifying the true seafloor reflection amidst the interfering OPS arrivals, a task for which conventional methods often prove inadequate. Furthermore, advanced techniques such as polarization filtering and GCDMO have limited applicability to legacy 2D datasets. Polarization filtering requires multi-component data, which is unavailable

in standard hydrophone streamer surveys. Similarly, deterministic GCDMO requires accurate cross-dip information, which is ill-posed to resolve using a single 2D line without 3D spatial sampling. These constraints highlight the need for a method that relies on morphological pattern recognition within the available scalar 2D data.

A further complication is that OPS can arise in different aspects of the seismic section. Depending on the geometry of the interfering structures, OPS may occur either before or after the true seafloor reflection. This overlap makes it particularly challenging in engineering and site-survey applications, where precise seafloor imaging is essential for placing infrastructure such as submarine cables or ocean-bottom instruments. By contrast, post-arrival OPS originates from deeper or more distant structures, where the signals mix with other lateral reflections and diffractions. These situations are more difficult to address with 2D methods and are often better suited to 3D imaging or advanced machine learning approaches that can use larger datasets. In this study, we concentrate on the pre-arrival case, which is both widespread and critical for practical applications because it directly overlaps with the water-bottom event. Our motivation arises from the fact that the first break picking can be used as a mute parameter to remove the OPS. This study uses and validates a semi-automated workflow that employs a lightweight, supervised neural network (NN) to accurately pick the true seafloor reflection, thereby enabling the muting of pre-arrival OPS. We apply this method to a 2D marine seismic line from the Ulleung Basin, in the East Sea, an area where irregular seafloor topography generates significant out-of-plane energy. This paper demonstrates that a targeted, picking-based workflow can effectively suppress OPS, improving the continuity and fidelity of the seafloor reflector in the final seismic image and offering a practical solution for enhancing the quality of legacy 2D seismic data.

## 2. Data and methodology

### 2.1. The study area and seismic data

The study area is located on the continental shelf and slope of the Ulleung Basin in the East Sea (Figure 1A).[20] The basin features water depths that generally increase from west to northeast. Notable seafloor topographic features create sudden changes in bathymetry (Figure 1B). To investigate structural and sedimentological variations, a 2D seismic survey was conducted in 2017. The seismic data were collected along an approximately 170 km line trending Northeast–Southwest (NE–SW) (shooting direction 225°). Key acquisition parameters are listed in Table 1. Near the northeastern part, adjacent to the start of the line, significant out-of-plane scattering originating from the irregular seafloor topography lateral to the survey track was observed (dashed box in Figure 1B and 1D). The complex 3D bathymetry of the survey area is further illustrated in the 3D views (Figure 1C). Figure 1D presents a zoomed 3D view of the dashed-box region alongside a brute-stack section, clearly showing false reflectors caused by the OPS.

These signals arrive earlier than the true seafloor reflection directly beneath the vessel. This occurs because the wavepath from the shallower, offline topographic features can be shorter than the path to the deeper seafloor directly below the line, even though the horizontal distance is greater. The workflow is specifically tailored for the common case where OPS from shallower, offline topography arrives before the true seafloor reflection. This pre-arrival OPS case is particularly challenging because it can lead to seafloor positioning errors for engineering applications. Conceptually, the two-way travel time for the true reflection ($t_{true}$) versus the OPS ($t_{OPS}$) can be compared:

$$t_{true} = \frac{2D_{vertical}}{V_{water}} \tag{1}$$

$$t_{OPS} = \frac{Slant\ path\ length}{V_{water}} \tag{2}$$

where, $D_{vertical}$ is the water depth directly beneath the survey line, and $V_{water}$ is the velocity of a seismic wave in water. Due to the geometry illustrated in Figure 2, it is often the case that $t_{OPS} < t_{true}$ when significant lateral topography exists at shallower depths than the seafloor below the vessel. Figure 2 provides a conceptual illustration showing how reflections from lateral seafloor topography can generate OPS that interferes with the desired reflection from the seafloor directly beneath the survey vessel in a 2D marine survey.

The entire brute-stacked section images the NE–SW trending Ulleung Basin, revealing shallow seafloor reflections at the ends of the line and a central slope region underlain by features interpreted as debris-flow and turbidite deposits (Figure 3A).[21,22] Figure 3B presents seven representative shot gathers sampled along the line.

An enlarged view of the target zone shows both relatively flat seafloor areas and the transition to the slope region (Figure 4A). Seismic trace headers usually contain recorded water bottom times. Figure 4B shows the result of applying a mute based on the water bottom time in the header information. However, this header-derived time is unreliable as it can be influenced by both the true seafloor reflection and strong OPS arrivals, as well as potential

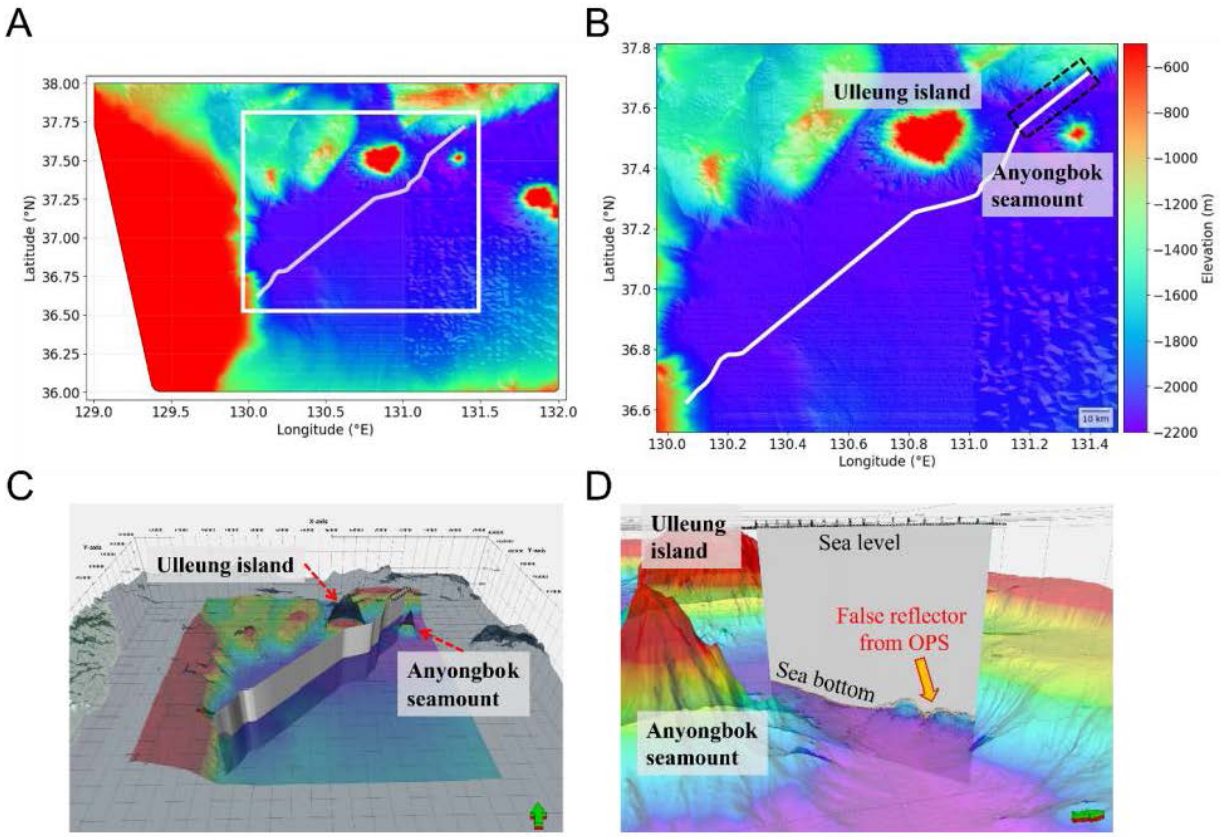



**Figure 1.** Overview of the survey area and geometry in the Ulleung Basin, East Sea. (A) Bathymetric map of the study region, with color shading indicating elevation above and depth below sea level; the white rectangle outlines the area enlarged in panel (B). (B) Zoom-in of the boxed region in (A) showing the vessel's survey track as a white line denoting the survey line; the dashed box highlights an area with significant out-of-plane signal (OPS); the scale bar in the lower right represents 10 km. (C) Three-dimensional perspective view of the survey area bathymetry. (D) Zoom-in three-dimensional perspective view of the dashed box region in (B), illustrating the seafloor topography causing OPS.

**Table 1. Key acquisition parameters for the two-dimensional marine seismic survey in the Ulleung Basin**

| Parameter | Value |
|---|---|
| Number of channels | 456 |
| Group interval (m) | 12.5 |
| Sampling interval (ms) | 1 |
| Recording length (s) | 7.0 |
| Air-gun volume ($in^3$) | 2,289 |
| Shot interval (m) | 25.0 |
| Shooting direction (degree) | 225 |

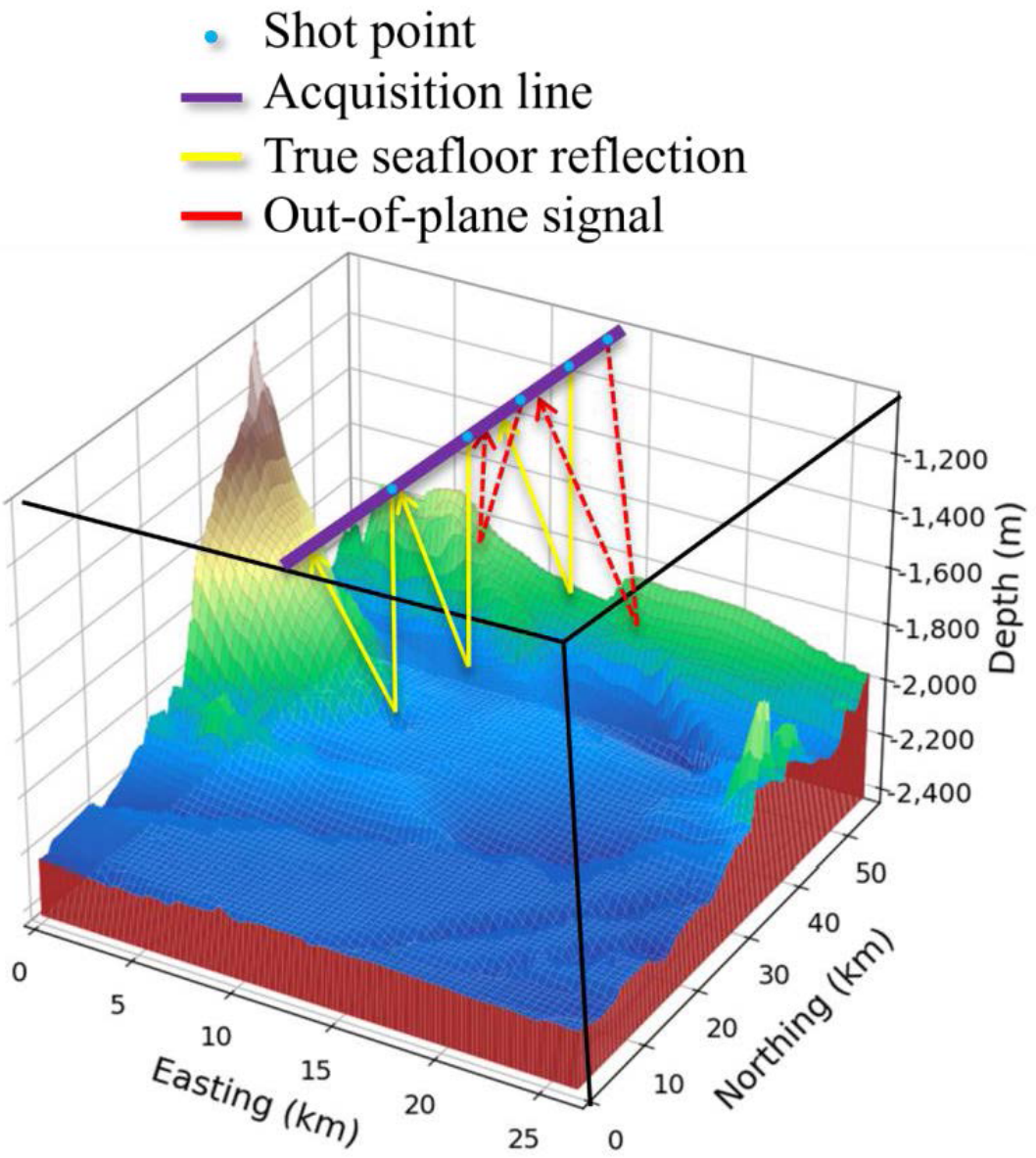



**Figure 2.** Conceptual illustration (pseudo–three-dimensional perspective) showing the generation of out-of-plane signal in a two-dimensional marine survey due to complex seafloor topography. The seismic source emits energy that reflects from the true seafloor directly beneath the vessel (solid yellow raypath, "true seafloor reflection") and also reflects off lateral topographic features outside the survey line (dashed red raypaths, "out-of-plane signal") arriving at the receiver and potentially interfering with the true reflection.

noise from instrument errors or weather conditions (dashed green arrows). Consequently, the mute based solely on the water bottom time in the header does not effectively remove the OPS and clarify the true seafloor. This result indicates that a simple, header-based approach is insufficient for accurate OPS removal in this dataset, motivating the evaluation of more advanced processing techniques.

### 2.2. Limitations of the conventional method: frequency–wavenumber filtering

As a baseline for comparison, we first assess the effectiveness of F–K directional filtering, a conventional method for attenuating coherent noise. We applied the filter independently to the same three domains used in our study, namely common shot gathers, common receiver gathers, and common depth point (CDP)–offset gathers.

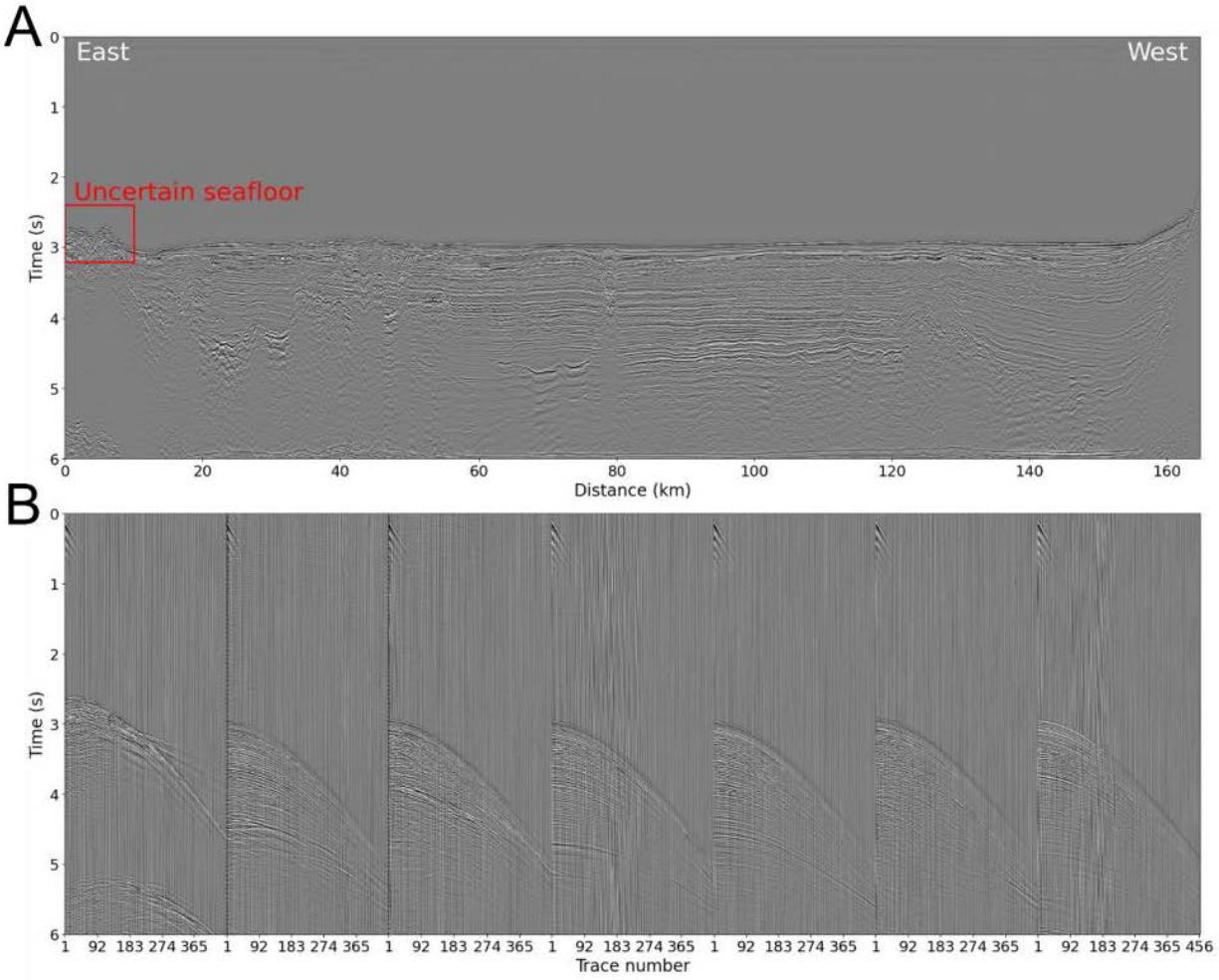



**Figure 3.** (A) Brute-stacked seismic section of the full line, with two-way travel time on the vertical axis and inline distance on the horizontal; the red box highlights the zone where the seafloor reflector is masked by out-of-plane signal. (B) Seven representative shot gathers were extracted at equal intervals along the line.

Figure 5 presents the F–K spectra before and after filtering for a representative gather from each domain.

The F–K spectra demonstrate that OPS energy exhibits an apparent moveout (dip) nearly identical to that of the true seafloor reflection. Consequently, their energy overlaps in the F–K domain, which prevents effective dip-selective filtering. Even with aggressive filtering applied, OPS events remained largely within the sections. The stacked sections in Figure 6 further confirm that OPS energy was only weakly suppressed and remained clearly visible after filtering. The failure of this conventional dip-based approach underscores the need for an alternative method capable of distinguishing OPS from true seafloor reflections using a broader combination of waveform attributes beyond just dip.

### 2.3. Neural network-based seafloor picking and muting

To overcome the limitations of conventional filtering, we use a technique centered on accurately picking the true seafloor reflection event and muting the preceding OPS arrivals. Our approach employs a supervised NN to learn the characteristics of the desired seafloor event from user-provided examples. Early studies demonstrated that NNs could automate seismic interpretation tasks, such as first-break picking by learning directly from data.[23,24] While modern deep learning techniques such as convolutional neural networks have shown success, they often require large labeled datasets,[25,26] which are rarely available for legacy 2D data with specific noise problems, such as OPS. Recent advances in machine learning have shown great

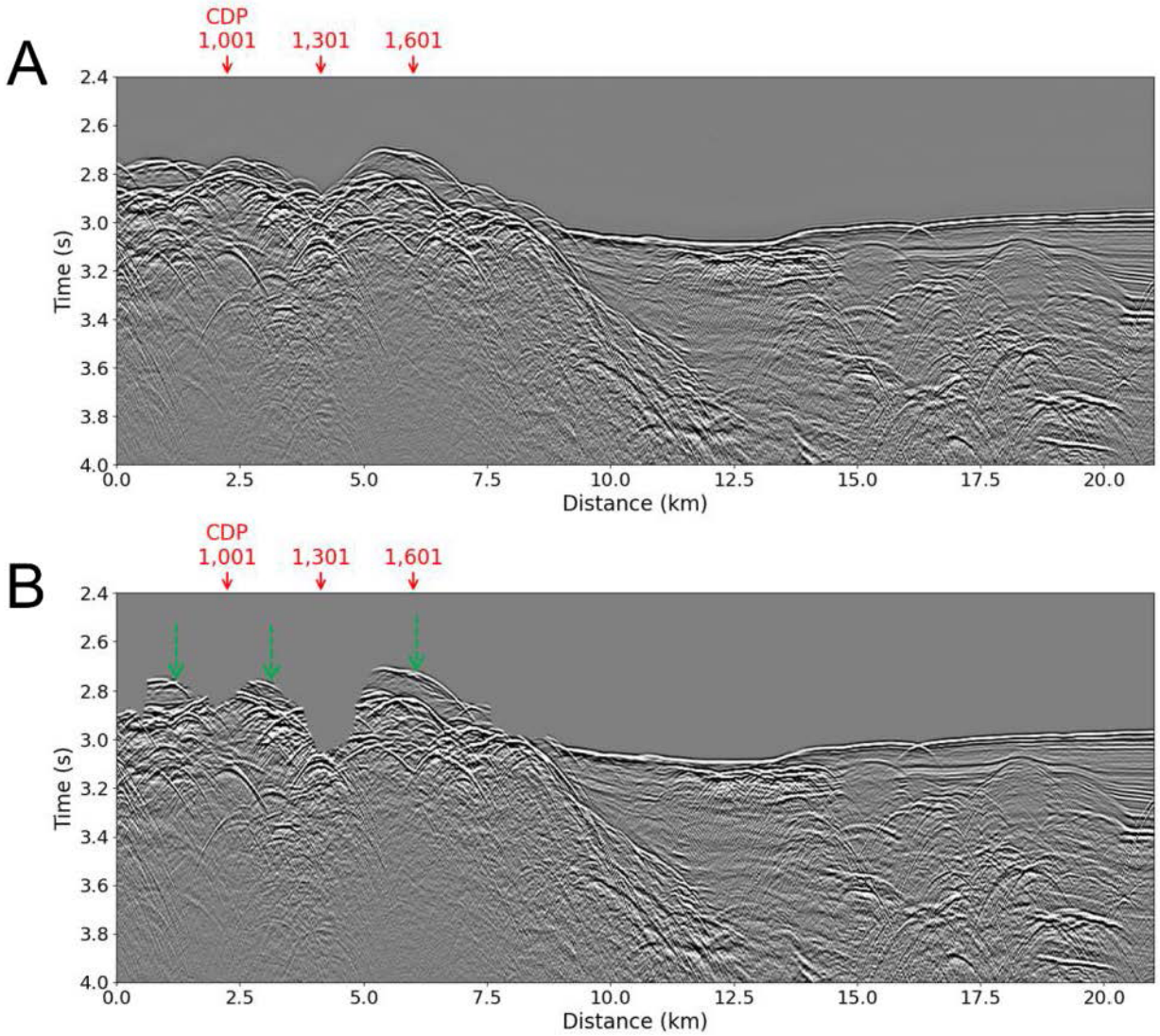



**Figure 4.** Enlarged view of the seafloor zone affected by out-of-plane signal (OPS) (0–21 km inline, 2.4–4.0 s two-way time), illustrating the OPS event and the result of header-based muting. (A) The brute-stacked section shows a coherent OPS event appearing near the seafloor between ~2.8 s and 3.2 s; solid red arrows mark common depth point (CDP) locations 1,001, 1,301, and 1,601. (B) The same interval after applying a mute based on the water bottom time in the header information. Note the persistence of residual OPS where the header mute does not accurately follow the true seafloor (dashed green arrow).

success in seismic processing, including robust automated picking across various environments[27,28] and sophisticated coherent noise attenuation using deep learning.[29,30] However, many of these deep learning denoisers[31] require large, high-quality labeled datasets, which are often unavailable for legacy 2D data that suffer from specific issues such as OPS. This constraint motivates the use of a more data-efficient strategy. Therefore, for this task, we selected the cascade–correlation architecture, a lightweight NN approach known for its ability to learn effectively from a small number of user-provided examples.[32,33] This makes it a practical choice for our specific goal of suppressing OPS in legacy data without requiring extensive pre-labeled datasets. The process involves defining a time window, manually picking the true seafloor event on a few traces, training the NN on these picks, and then applying the trained network to pick the event automatically on all traces within the target zone (Figure 7). The neural network inputs consist of seismic trace amplitudes extracted from a user-defined time window centered on the manually picked event. This window, or "gate," allows the network to capture the local waveform shape and moveout characteristics of

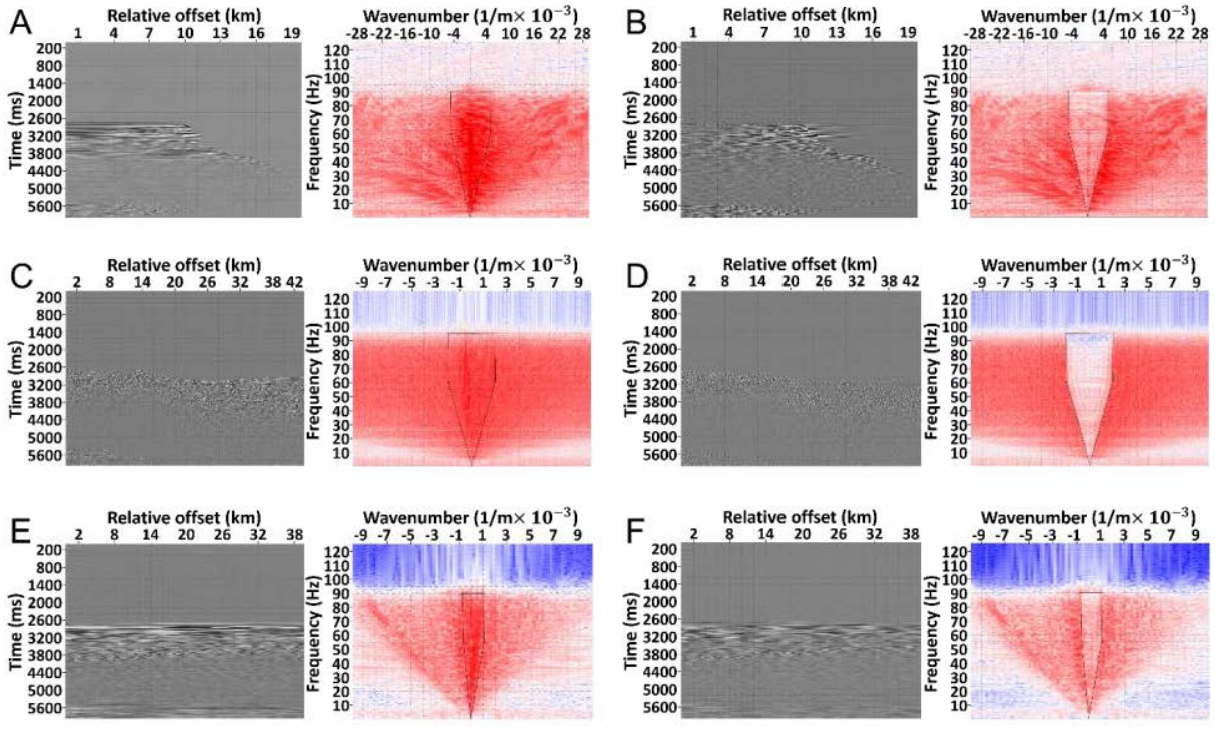



**Figure 5.** Representative gathers and their F–K spectra before and after F–K directional filtering. (A, B) Common shot gather. (C, D) Common receiver gathers. (E, F) Common depth point–offset gather. Note the spectral overlap between the seafloor reflection and the out-of-plane signal energy, which limits the effectiveness of dip-selective filtering.

the true seafloor reflection. For this study, the gate was set to encompass approximately two periods of the dominant frequency around the first break. The network minimizes the error between the manual pick time and its predicted arrival time by adjusting the weights of the added hidden units. The specific training parameters used are detailed in Table 2. Arrival time was predicted by fitting a line to several previous picks using the slope of the top gates. Thus, the window in which the network considers a pick can be restricted to a small range about the predicted arrival time of the first break. Once the NN pick determines the arrival time of the true seafloor reflection, this horizon information is converted into a mute table, which is then used to apply a top mute, removing energy arriving before the picked time (i.e., the OPS). This workflow is designed to generate an accurate time horizon for the true seafloor that can be used for muting or other interpretation purposes.

### 2.4. The cascade–correlation architecture

To discriminate OPS from true reflections, we employed the NN picking tool available in Landmark's SeisSpace processing software. This tool utilizes a supervised artificial neural network based on the cascade–correlation architecture developed by Fahlman and Lebiere.[32] Unlike standard backpropagation networks, which adjust weights within a fixed architecture, cascade–correlation incrementally builds the network structure during training. It starts with a minimal network (inputs connected directly to outputs) and then iteratively adds new hidden units one by one. Crucially, once a hidden unit is added, its input weights are frozen. Only the weights for the new unit are trained, followed by retraining only the output layer weights. A new hidden unit $h_j$ receives connections from all input units $x_i$ and all previously installed hidden units $h_k$ (where $k < j$). Its output $y_j$ is calculated as

$$y_j = \sigma\left(\sum_i w_{ji} x_i + \sum_{k<j} w_{jk} y_k\right) \tag{3}$$

where $w$ are the input weights and $\sigma$ is an activation function (e.g., sigmoid). Each new hidden unit is trained to maximize the correlation between its output and the residual error signal that the existing network has not yet learned. Specifically, the algorithm attempts to maximize a measure $S$, the magnitude of the covariance between the candidate unit's output, $V$, and the residual output error,

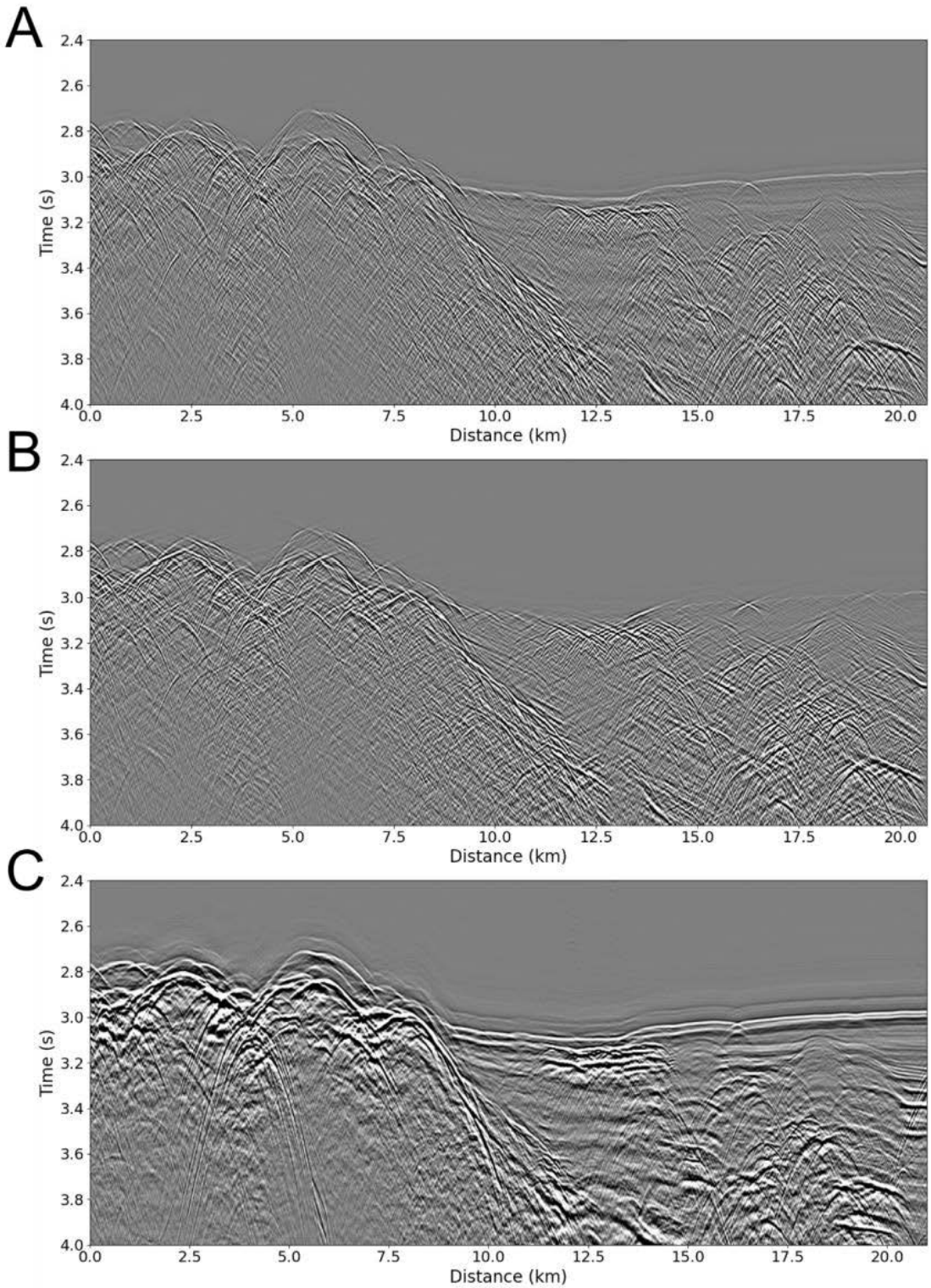



**Figure 6.** Stacked sections after F–K directional filtering were applied in the (A) common shot, (B) common receiver, and (C) common depth point-offset domains. Residual out-of-plane signal remains visible near the seafloor in all cases.

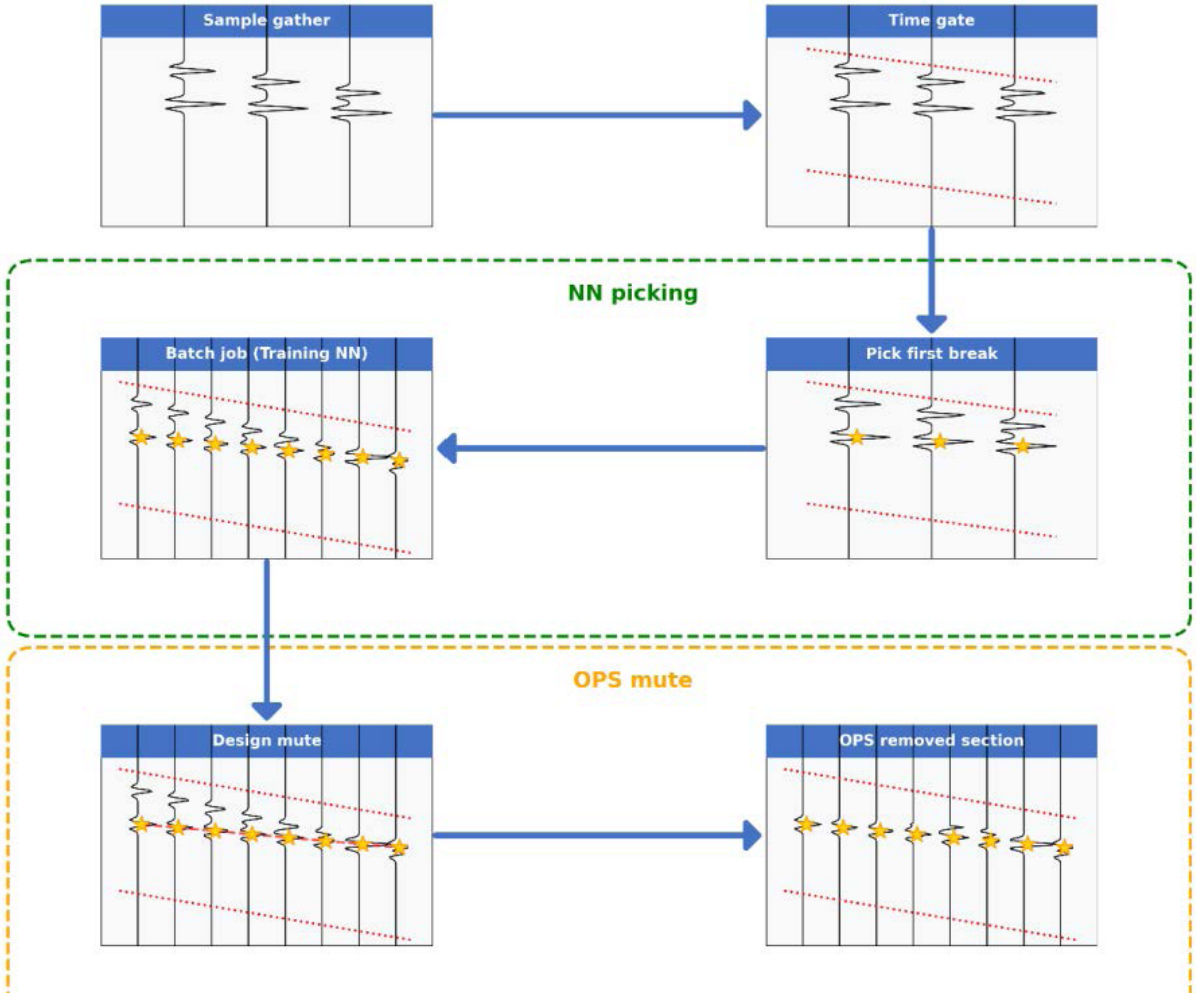



**Figure 7.** Processing flow to determine the mute table from the neural network (NN) pick for out-of-plane signal (OPS) removal

**Table 2. Parameters in neural network for first-break (FB) picking in seismic data processing**

| Parameter | Value/description |
|---|---|
| Number of traces in the median line fit | 6 |
| Maximum trace-to-trace static (ms) | 20.0 |
| Starting offset to determine FB pick slope (m) | 150.0 |
| Time gate file/database name | Filepath & name |

$E_o$:

$$S = \sum_o \left| \sum_p \left( V_p - \bar{V} \right) \left( E_{p,o} - \bar{E}_o \right) \right| \quad (4)$$

where the sum is over output units $o$ and input patterns $p$, and $\bar{V}$ and $\bar{E}_o$ represent average values.

This incremental learning strategy, focusing on new units on the remaining error, allows the network to adapt its complexity to the specific problem. This approach was chosen for its efficiency and demonstrated ability to train effectively on relatively few user-provided samples. A schematic illustration of this architecture is shown in Figure 8.

### 2.5. Training process and feature recognition

The training process is an iterative, gather-by-gather workflow. First, an interpreter manually picks the entire seafloor reflection horizon on a single representative gather, which serves as the label data. The network is trained using this single, complete example and then immediately evaluated by having it predict the horizon on the very same gather. If the prediction does not closely match the manual pick, the training process is repeated to refine its parameters. This iterative training-and-evaluation loop continues until the network can accurately reproduce the manual horizon for that gather. Once the NN picking is complete, the determined travel times are saved to the seismic database and transferred to the trace headers as the seafloor arrival time. This horizon information is then directly mapped to the mute parameter header to define the top-mute function. No additional tapering or smoothing was applied during this transfer, preserving the exact pick times determined by the network.

This process is illustrated in Figure 9. The entire iterative cycle is then performed on approximately 20 different representative gathers, chosen to encompass variations in water depth and OPS intensity, to build a robust final network. The key configuration parameters of the network training are summarized in Table 3.

The primary physical features that the NN is designed to differentiate can be grouped into several categories. Residual moveout after normal moveout (NMO) correction provides the most important feature, since the true in-plane seafloor reflection becomes relatively flat while OPS retains curvature. Additionally, lateral continuity and coherence help the network identify the true seafloor, which generally exhibits greater spatial continuity than OPS events that are more discontinuous. Finally, waveform and peak travel time provide an additional constraint, as the implementation was designed to track the travel time of the reflection's peak amplitude, linking this with the other physical properties of the true seafloor. As shown in Figure 9, the network progressively refines its ability to combine these indicators, aligning its predictions more closely with the interpreter-defined horizon through repeated adjustment.

### 2.6. Application in different gathering domains

The entire workflow was executed in three different gather domains to evaluate which domain yielded the best performance for OPS removal, including common shot gathers (conventional shot records in streamer surveys), common receiver gathers (grouping traces by receiver channel across all shots), and CDP–offset gathers (conventional common midpoint gathers). A separate and independent training process was carried out for each data domain. For the common shot gathers, a distinct network was trained using manual picks from that domain. The same procedure was then repeated for the common receiver and CDP–offset gathers, ensuring that each domain produced a network specifically optimized for its own data representation. Applying NMO correction to common shot gathers can help flatten the primary seafloor reflection, which may facilitate NN picking and subsequent OPS muting in this domain. Conversely, common receiver gathers typically exhibit less severe moveout for the seafloor event, often eliminating the need for NMO correction before picking, which simplifies the process and avoids potential NMO stretch artifacts. We evaluated the effectiveness of NN picking in each domain.

The network learns to distinguish the seafloor event from the OPS based on characteristics learned from the training examples, such as amplitude, frequency content, waveform shape, and continuity across traces. For instance, the OPS might have different amplitude decay or moveout characteristics compared to the true seafloor reflection originating from directly beneath the line. Because the network is trained locally on the specific survey data, it adapts to the particular noise and signal characteristics present. A key aspect of this technique is that it does not require a large labeled dataset or extensive training time like many modern deep learning approaches. The NN essentially acts as an intelligent tool to propagate the interpreter's identification of the target event across the dataset based on learned patterns.

## 3. Results

### 3.1. Application of neural network picking to seismic gathers

We applied the NN picking workflow in the three gather domains (common shot, common receiver, and CDP–

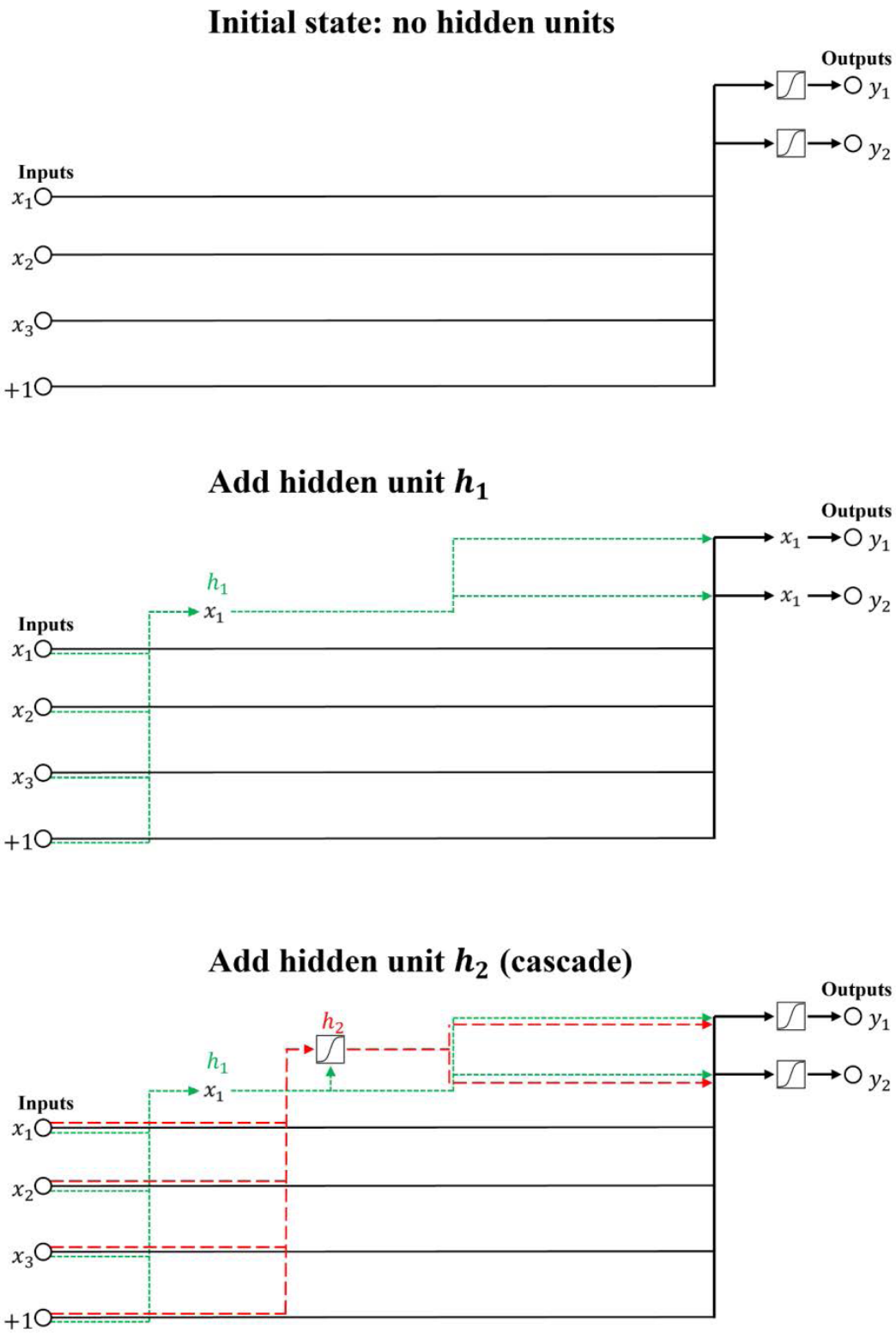



**Figure 8.** Schematic diagram of the cascade–correlation neural network architecture (adapted from Fahlman and Lebiere[32]). The network incrementally adds hidden units ( $h_1$ , $h_2$ , etc.), which receive input from all previous input and hidden units. Once a unit is added, its input weights are frozen.

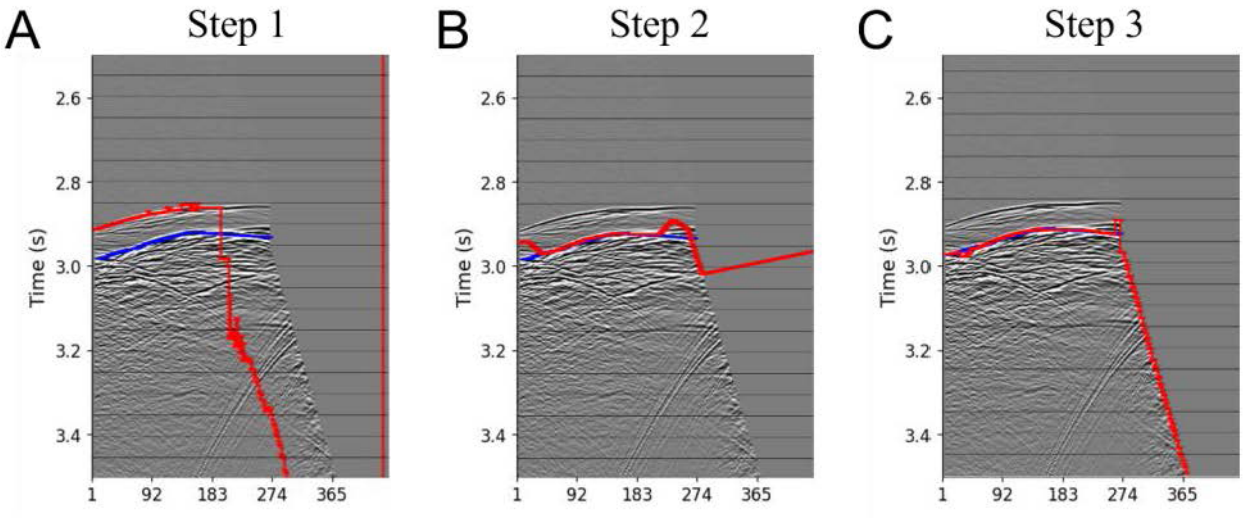



**Figure 9.** An iterative training process for the neural network on a representative dataset. The blue line is the manual pick (training label), and the red line is the neural network prediction. (A) After the first pass, the prediction is inaccurate. (B) After the second pass, the prediction improves. (C) After the third pass, the prediction converges with the manual pick.

**Table 3. Summary of key parameters for the neural network training**

| Parameter | Value/description |
|---|---|
| Neural network algorithm | Cascade–correlation |
| Software implementation | Landmark SeisSpace (Neural Network First Break Picker module) |
| Training strategy | Iterative, gather-by-gather refinement |
| Number of training gathers | Approximately 20 representative gathers |
| Labeling method | Manual picking of the complete seafloor horizon on each training gather |
| Target feature | Peak-amplitude travel time of the seafloor reflection and the manual pick |
| Stopping criterion | Visual confirmation of convergence between the network's prediction |
| Final application | The fully trained network is applied in batch mode to the entire dataset |

offset) to assess the performance of OPS removal in each case. Figures 10, 11, and 12 illustrate representative gathers for each domain, showing the data before muting, the determined NN pick, and the gathers after muting for each domain, respectively. The gathers shown are extracted from locations near CDPs 1,001, 1,301, and 1,601, which fall within the zone affected by OPS (indicated by solid arrows in Figure 4A).

For the common shot gathers, Figure 10A displays sample NMO–corrected common shot gathers. The NMO correction adjusts the travel times based on the source–receiver offset ($x$) according to the hyperbolic travel time equation:

$$t(x)^2 = t(0)^2 + \frac{x^2}{V_{NMO}^2} \tag{5}$$

where, $t(x)$ is the two-way reflection time at offset $x$, $t(0)$ is the two-way zero-offset time, and $V_{NMO}$ is the NMO velocity. In the enlarged NMO-corrected view (Figure 10B), the primary seafloor reflection events should ideally be aligned horizontally after correction using the appropriate velocity. In contrast, the OPS, originating from offline sources with different travel paths, does not align horizontally, which helps in its visual discrimination. The red line overlaid on the gathers in Figure 10B indicates the horizon picked by the NN, targeting the true seafloor reflection. Determining the correct mute time manually is challenging due to the interfering OPS and variations in seafloor topography along the survey line. After applying the mute based on the NN picks (Figure 10C), the OPS is mostly removed. However, some residual noise remains, and potential over-muting (clipping the start of the seafloor reflection) is observed in some areas, reflecting the challenges of picking accurately in this domain.

Figure 11A shows sample common receiver gathers where the OPS is visible ahead of the true seafloor reflection. Because these gathers exhibit less moveout delay for the seafloor event compared to shot gathers, NMO correction was not applied before picking. The red line in Figure 11B shows the NN pick for the true seafloor. The resulting mute (Figure 11C) effectively removes the OPS, allowing the true seafloor reflection to be observed more clearly and continuously. The effectiveness in this domain benefits from the relatively long streamer configuration used in the survey, which provides sufficient data redundancy for the NN to learn the patterns.

Finally, for the CDP–offset gathers, OPS can be observed before the stacking result (Figure 12A). After NMO correction, both the OPS and the true seafloor reflection appear relatively flat because their travel paths are mostly through the water column. The NN pick (red line in Figure 12B) targets the later-arriving true seafloor event. Applying the mute based on these picks removes the OPS (Figure 12C). However, similar to the common shot results, some inaccuracies lead to residual noise or over-muting, especially at the far offset traces. While velocity analysis is typically performed in the CDP–offset domain, which could make OPS removal here convenient, the large number of traces and potential smearing effects during sorting may have made the NN training and picking less consistent compared to the common receiver domain for this dataset.

### 3.2. Quantitative comparison of picking consistency

Before stacking the data, we performed a statistical analysis of the NN-picked travel times to quantitatively assess the consistency of the results from each domain (Table 4). The analysis indicates that picking in the common receiver gather domain yielded the most consistent results, as evidenced by the lower standard deviation (67.5 ms) compared to common shot (129.7 ms), and CDP–offset gathers (98.4 ms). This suggests that the cascade–correlation neural network's ability to generalize the trend of the true seafloor reflection was most effective in the common receiver domain, likely due to the more consistent presentation of both the true seafloor event and the OPS pattern in this domain for this dataset.

### 3.3. Analysis of stacked and migrated sections

Before analyzing the final stacked sections, the consistency of the picks from the three domains was visually inspected. The picks were transferred and plotted onto a single NMO–corrected CDP gather (CDP 1,401) to directly compare their alignment and flatness, as shown in Figure 13. Note that the lateral coverage of the picks in Figure 13A–C may vary slightly, as the original picking was performed in domains (shot, receiver, and CDP) encompassing different sets of traces relative to CDP 1,401. This comparison provides a visual basis for the quality of the subsequent stacks.

Figure 14 compares the stacked sections after applying the NN pick mute derived from the different gather domains. Muting based on common shot gather picks removed much of the OPS between 2,600 and 3,000 ms, but some residual energy and discontinuities remain along the seafloor reflector, likely due to inconsistencies in the picks from this domain (Figure 14A). Beneath the seafloor event, numerous diffraction hyperbolas are visible, likely originating from irregular subsurface geological structures. On the other hand, muting based on common receiver gather picks yielded the most consistent result (Figure 14B). The quality of the mute appears more consistent,

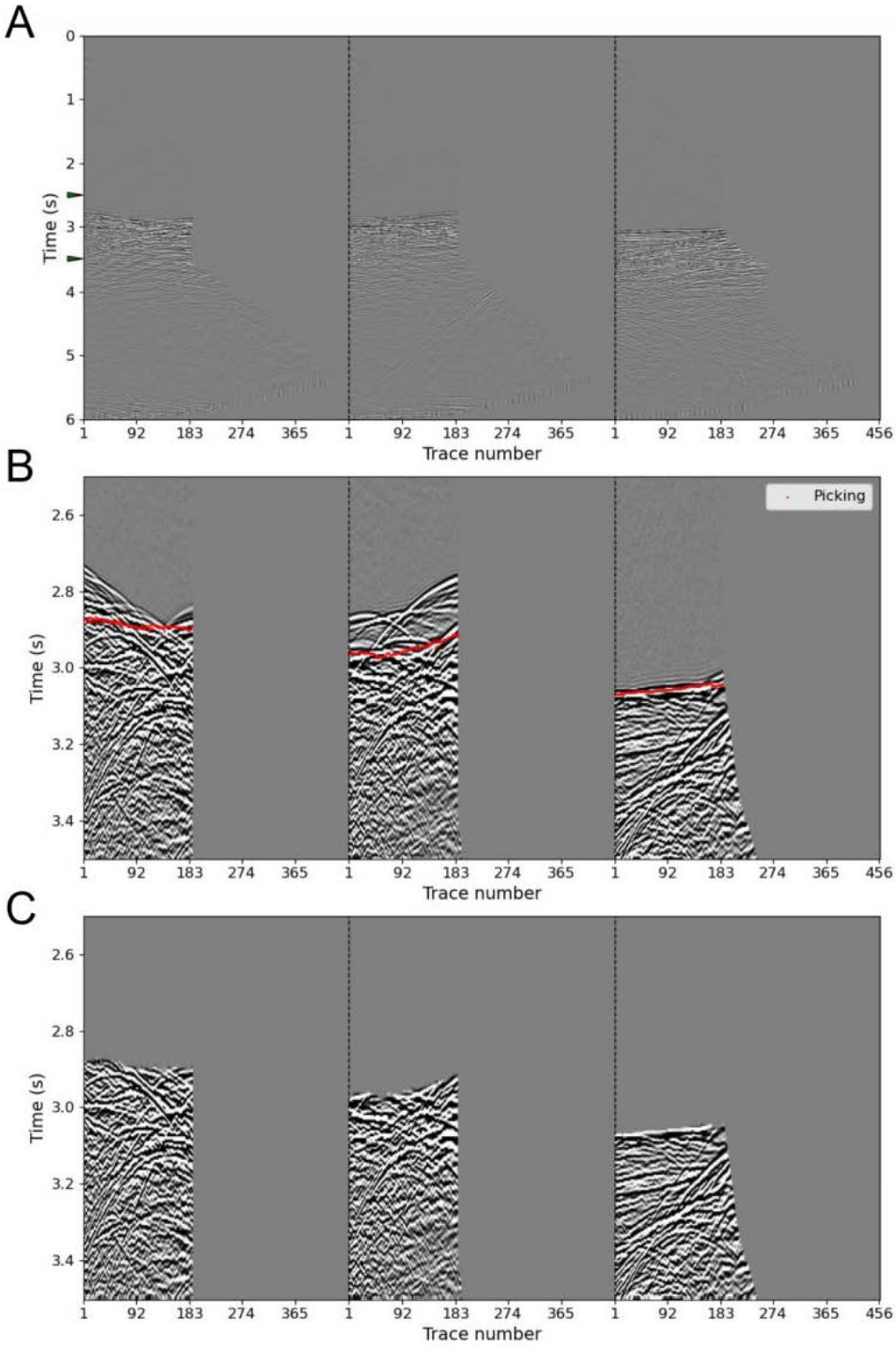



**Figure 10.** Out-of-plane signal removal workflow in the normal moveout–corrected common shot gather domain at three locations (near common depth points 1,001, 1,301, 1,601). (A) Gathers before muting, showing out-of-plane signal. (B) Neural network pick (red line) targeting the true seafloor reflection. (C) Gathers after applying the neural network-derived mute.

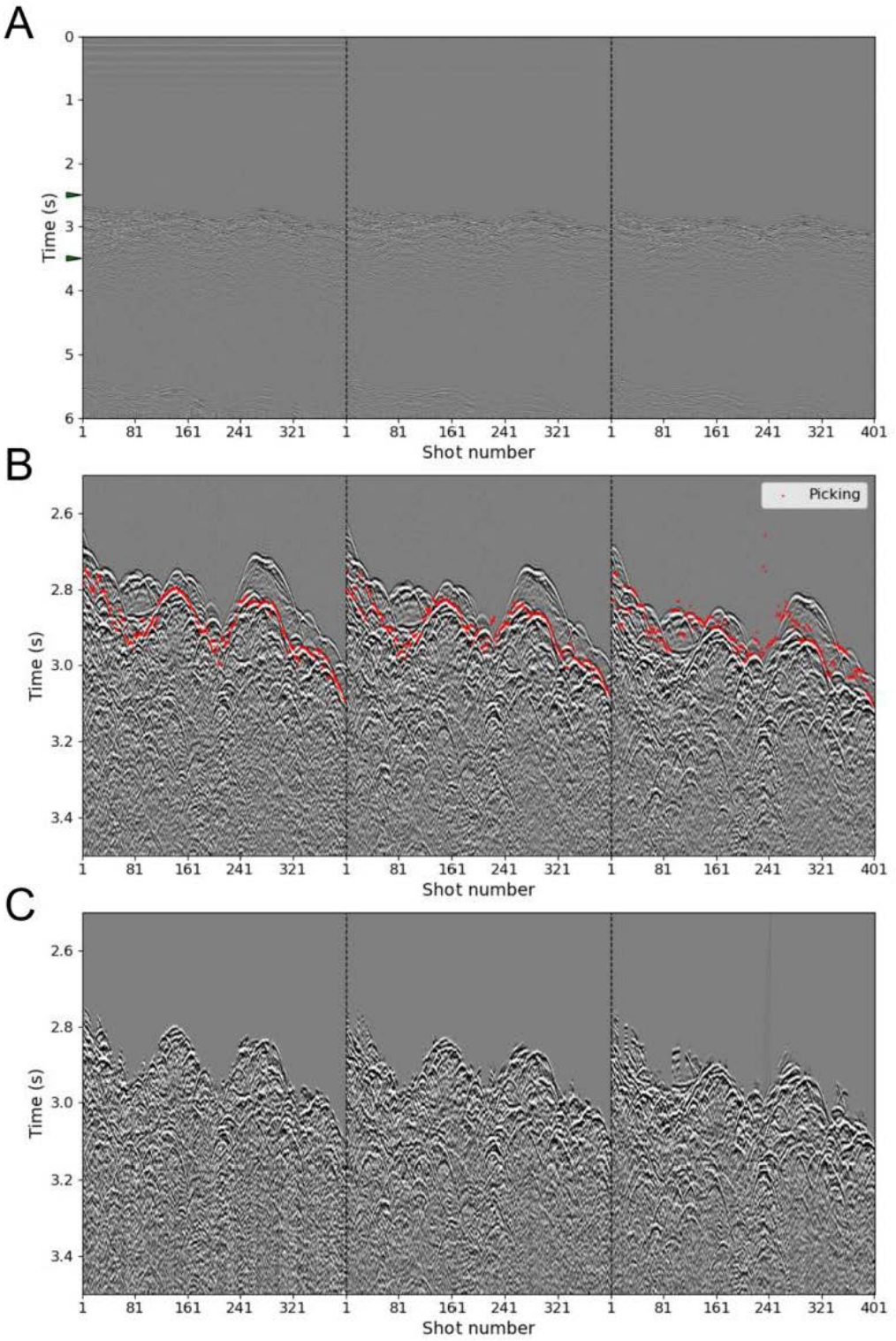



**Figure 11.** Out-of-plane signal removal workflow in the common receiver gather domain at three locations (near common depth points 1,001, 1,301, 1,601). (A) Gathers before muting, showing out-of-plane signal. (B) Neural network pick (red line) targeting the true seafloor reflection. (C) Gathers after applying the neural network-derived mute.

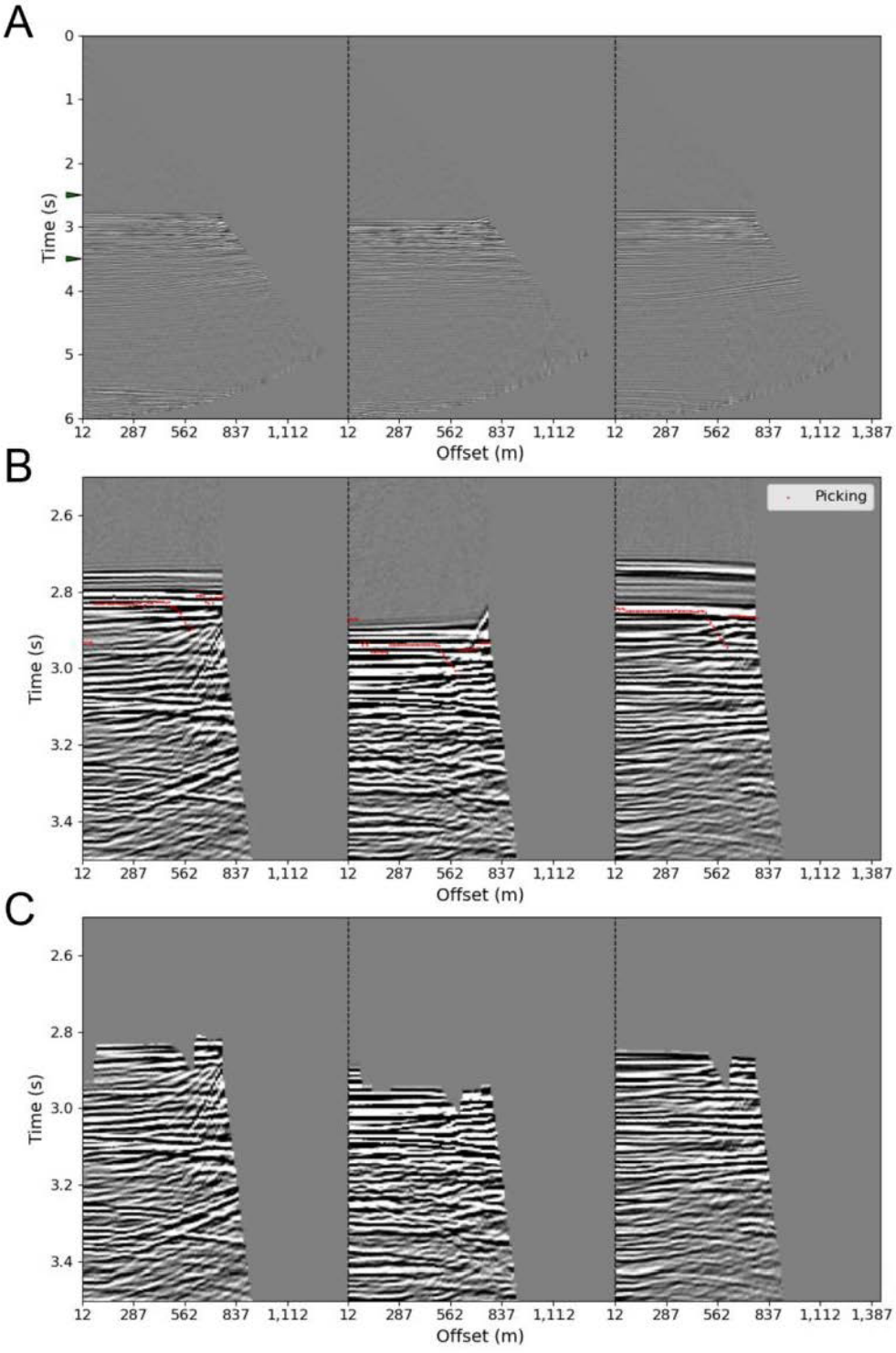



**Figure 12.** Out-of-plane signal removal workflow in the normal moveout–corrected common depth point–offset gather domain at three locations (near common depth points 1,001, 1,301, 1,601). (A) Gathers before muting, showing out-of-plane signal. (B) Neural network pick (red line) targeting the true seafloor reflection. (C) Gathers after applying the neural network-derived mute.

**Table 4. Statistics of neural network picking results**

| Statistic | Common shot gather pick | Common receiver gather pick | Common depth point–offset gather pick |
|---|---|---|---|
| Minimum time | 2,518.9 | 2,688.6 | 2,534.3 |
| Maximum time | 3,206.2 | 3,108.8 | 3,131.4 |
| Mean | 2,964.2 | 3,043.6 | 2,828.7 |
| Standard deviation | 129.7 | 67.5 | 98.4 |

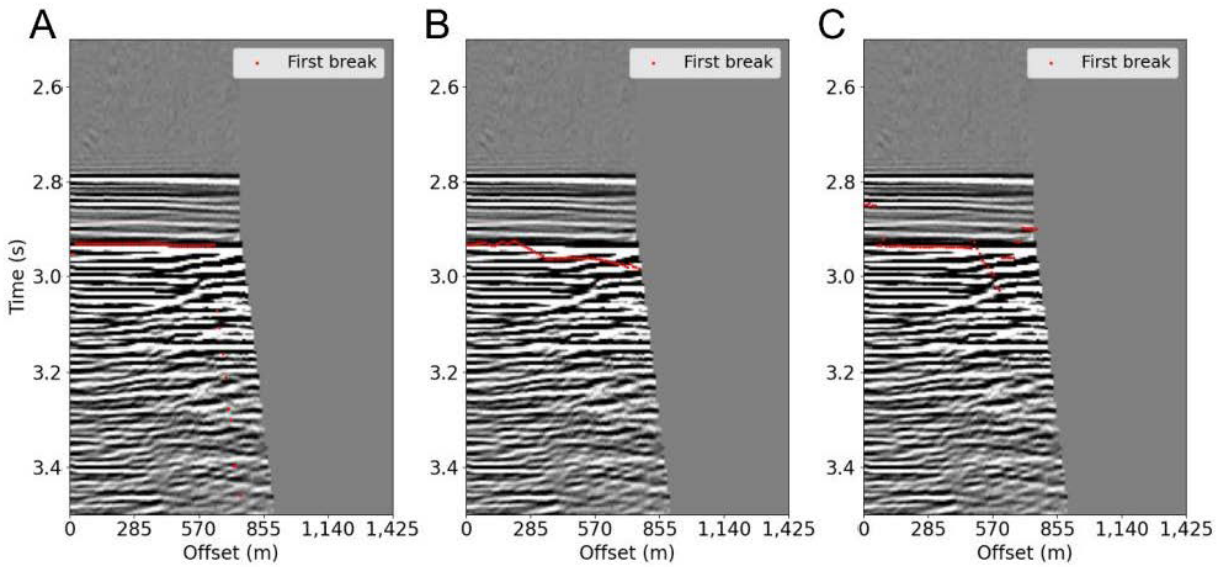



**Figure 13.** Comparison of neural network picking from the three domains, plotted on the normal moveout–corrected common depth point (CDP) gather at CDP 1,401. The picks were derived from the (A) common shot, (B) common receiver, and (C) CDP–offset domains.

resulting in improved continuity of the seafloor event compared to the other domains. Muting based on CDP–offset gather picks showed partial removal of OPS, but the result is similar to the common shot case, with some remaining discontinuities (Figure 14C). The challenges in training the NN effectively in this domain, possibly due to the large number of traces and sorting effects, likely contributed to a less effective outcome.

To collapse diffractions and accurately position the reflections, we applied Kirchhoff post-stack time migration to the muted datasets (Figure 15). This process sums seismic amplitudes along diffraction curves defined by the migration velocity ($V_{mig}$). The basic principle can be represented by the Kirchhoff summation or integral:[34,35]

$$P_{mig}\left(x_{out},t_{out}\right)\approx\sum_{x_{in}}\omega\left(x_{in},x_{out},t_{out}\right)P_{stack}\left(x_{in},t_{in}\left(x_{in},x_{out},t_{out}\right)\right) \quad (6)$$

where $P_{mig}\left(x_{out},t_{out}\right)$ is the migrated amplitude at the output location $x_{out}$ and zero-offset time $t_{out}$, $P_{stack}\left(x_{in},t_{in}\right)$ is the stacked amplitude at the input location $x_{in}$ and input time $t_{in}=\sqrt{t_{out}^2+\frac{4\left(x_{in}-x_{out}\right)^2}{V_{mig}^2}}$ (which lies on the diffraction

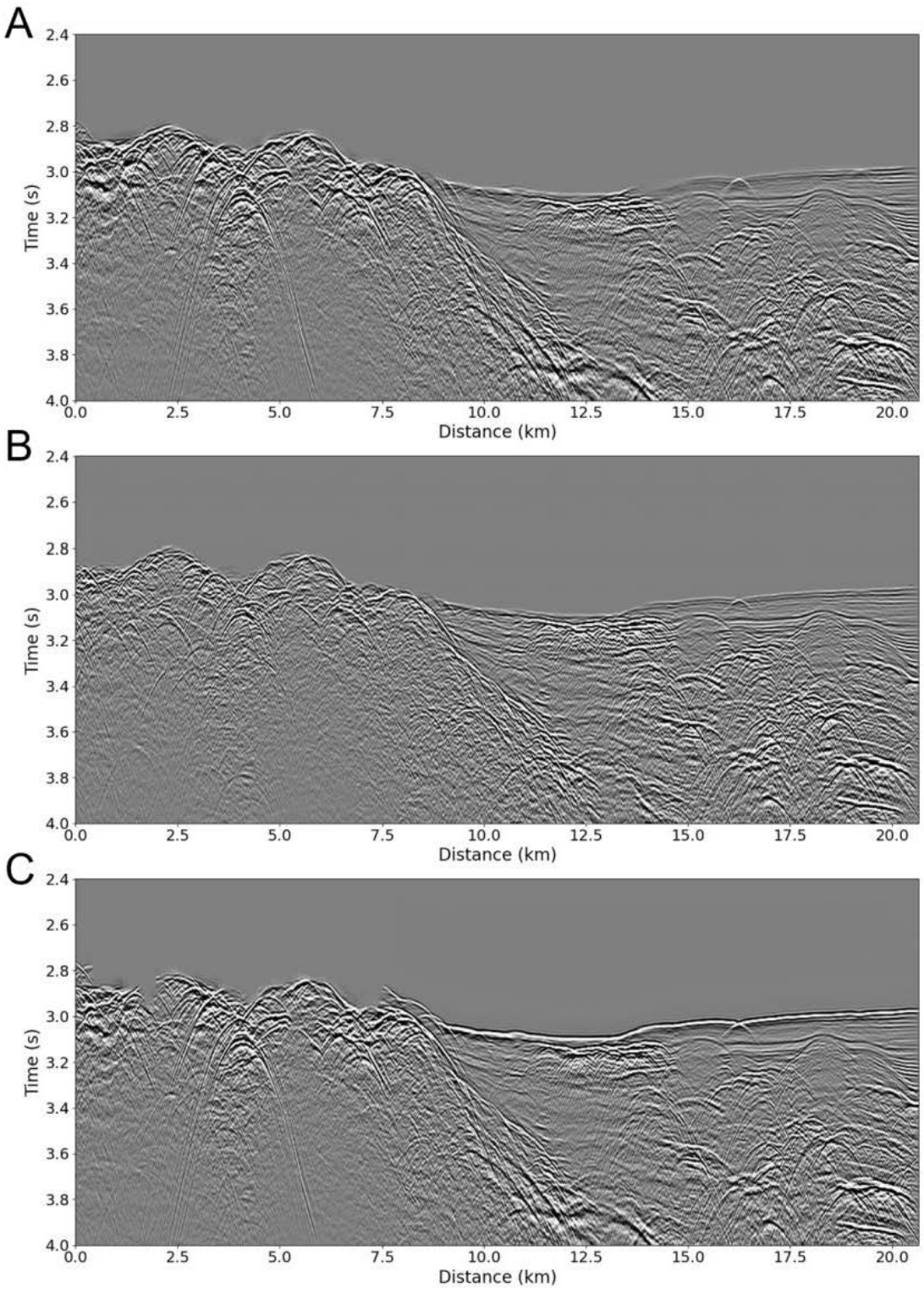



**Figure 14.** Stacked sections after applying the neural network-derived mute in three different domains. (A) Muting in the common shot domain. (B) Muting in the common receiver domain shows the most continuous seafloor reflection. (C) Muting in the common depth point–offset domain.

curve), and $\omega$ represents appropriate weighting factors (e.g., for amplitude scaling and filtering). Migration improved the image of the seafloor in all cases compared to the stacks. The migrated sections show the correction of diffraction signals originating from below the seafloor, resulting in improved overall image quality compared to the brute stack section. Consistent with the stacked results, the migration of the common receiver gather muted data (Figure 15B) exhibits a cleaner and more continuous seafloor reflection, appearing smoother than the results from the common shot (Figure 15A) and CDP–offset (Figure 15C) muted data. The common shot (Figure 15A) and CDP–offset result (Figure 15C) show improvement over the stack but still appear slightly less continuous than the common receiver result. However, where the OPS and true seafloor reflection overlap significantly in time, full recovery of the primary reflection remains difficult even after muting, as the muting process removes energy but cannot restore the signal masked by the overlap.

### 3.4. Validation against external data

To move beyond qualitative inspection, we quantitatively validated the seafloor picking results using an independent bathymetric reference from the Korea Hydrographic and Oceanographic Agency (KHOA).[20] The KHOA bathymetric profile along the survey line was converted to two-way travel time using a water velocity of 1,480 m/s and compared with the seismic seafloor picks. We defined a travel time-delay metric to evaluate accuracy, computing both the mean absolute error (MAE) and the standard deviation of the time difference at each CDP location. Figure 16 illustrates the metric, where Figure 16A shows the brute stack (identical to the original Figure 4A) with the KHOA seafloor profile (solid line) and, for context, the seafloor inferred from first-arrival reflection (dashed line). Figure 16B shows the header-muted stack (Figure 4B) with the same overlays.

The results, summarized in Table 5, show improvement after applying NN-based OPS suppression. The unprocessed brute stack exhibited an MAE of about 115 ms with a standard deviation of ~33 ms, indicating large discrepancies relative to the true seafloor. A header mute reduced the MAE to ~71 ms, though variability remained high (standard deviation ~41 ms). By contrast, the NN-based workflow achieved lower errors where the stacked sections showed MAE values of approximately 45–61 ms, with reduced variability, and the migrated sections further improved to 47–50 ms MAE with standard deviations generally below 38 ms. In contrast, the F–K filtered stacks retained large errors (MAE ~120–216 ms), nearly as high as or worse than the brute stack, confirming that OPS energy remained and biased the seafloor picks. Assuming a water velocity of 1,480 m/s, an MAE of 10 ms in two-way travel time corresponds to a depth discrepancy of approximately 7.4 m. Consequently, the reduction in error achieved by the proposed method significantly improves the accuracy of seafloor depth estimation compared to the brute stack. These quantitative metrics demonstrate that NN-aided muting yields a seafloor horizon substantially more consistent with the independent bathymetric reference than either no OPS removal or conventional filtering approaches.

We also compared the seismic images directly with the external reference. Figure 17 overlays the KHOA bathymetry (cyan line) on the brute stack, the header-muted stack, and the final migrated sections. In the brute stack (Figure 17A), strong OPS obscures the true seafloor, resulting in large deviations between the cyan line and the interpreted horizon. With a header mute (Figure 17B), some OPS is attenuated, but the remaining residuals still distort the seafloor pick. In contrast, after applying the NN-based suppression, the final migrated sections (Figure 17C–E) show the reflector aligning much more closely with the true seafloor profile, with notably improved positional consistency in the 2.7–3.0 s window.

Finally, to verify that the coherent energy removed by the workflow corresponds to the OPS noise (and not a legitimate signal), we computed difference sections. Figure 18 shows the absolute difference between the brute stack (Figure 4A) and each final muted stack (Figure 14A–C), focused on a time window of 2.4–3.2 s and the first 8 km of the line. This process highlights the energy eliminated by the mute. The difference displays reveal banded coherent events where the OPS was present, indicating that those events were removed. Notably, in the 2.7–3.1 s range, the shape and location of the removed energy match the OPS noise seen in the original data. Some areas show very little difference (meaning the seafloor event was preserved), whereas areas of intense difference correspond exactly to where OPS interfered. This confirms that the technique targeted and removed the intended OPS signals. It is also worth noting that the CDP–offset gather case shows relatively weaker OPS suppression near 7.5–8 km, where the difference is less pronounced. In contrast, the shot and offset gather domains exhibit clearer energy removal in the same region. This suggests that the effectiveness of the technique can vary by data domain, with shot and offset gathers providing more effective suppression in this dataset.

## 4. Discussion

Seismic data acquired in areas of irregular seafloor morphology often contain OPS that can mask the true

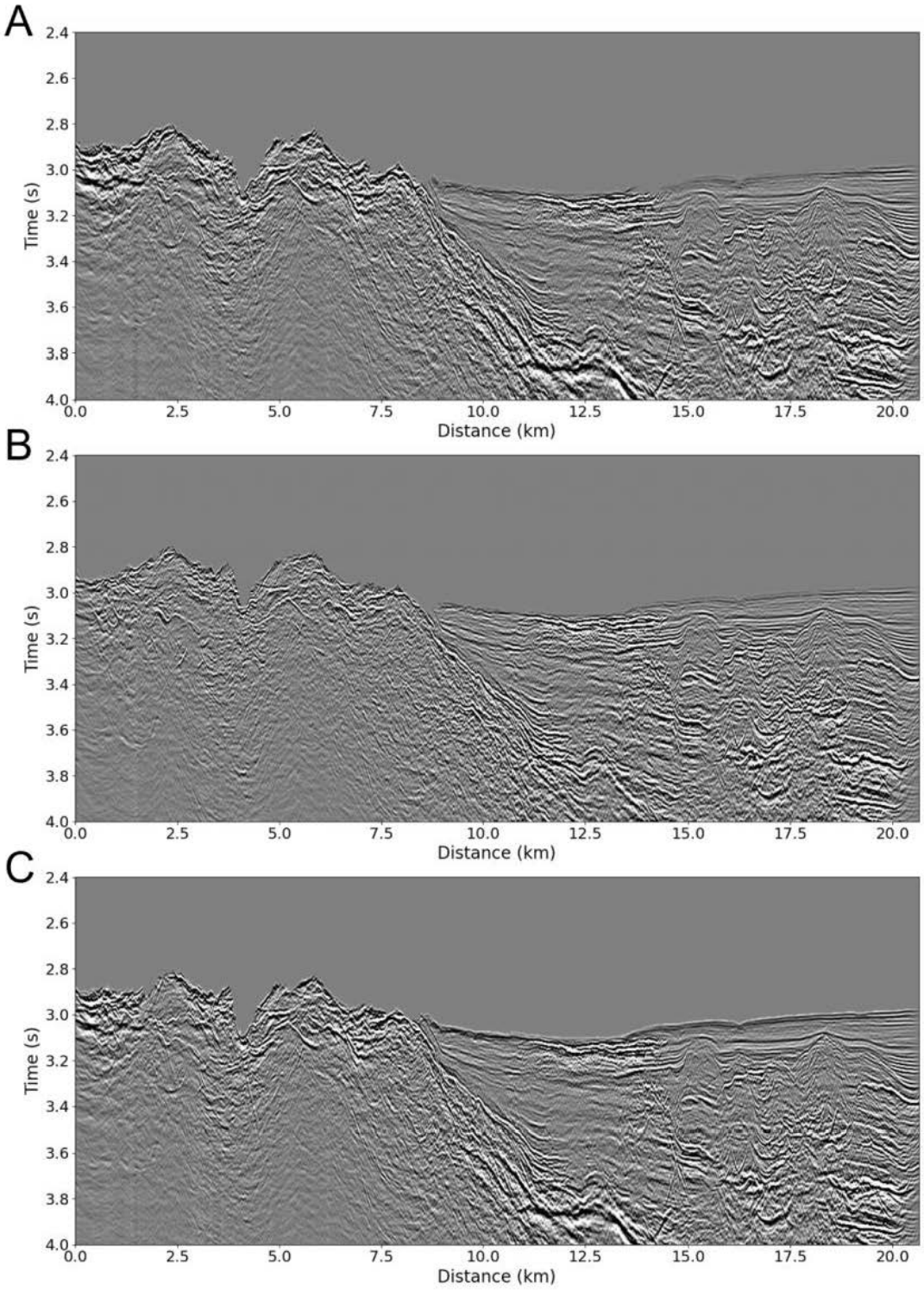



**Figure 15.** Kirchhoff post-stack time migration of the muted sections from Figure 14. (A) Migration of the common shot muted data. (B) Migration of the common receiver muted data. (C) Migration of the common depth point–offset muted data.

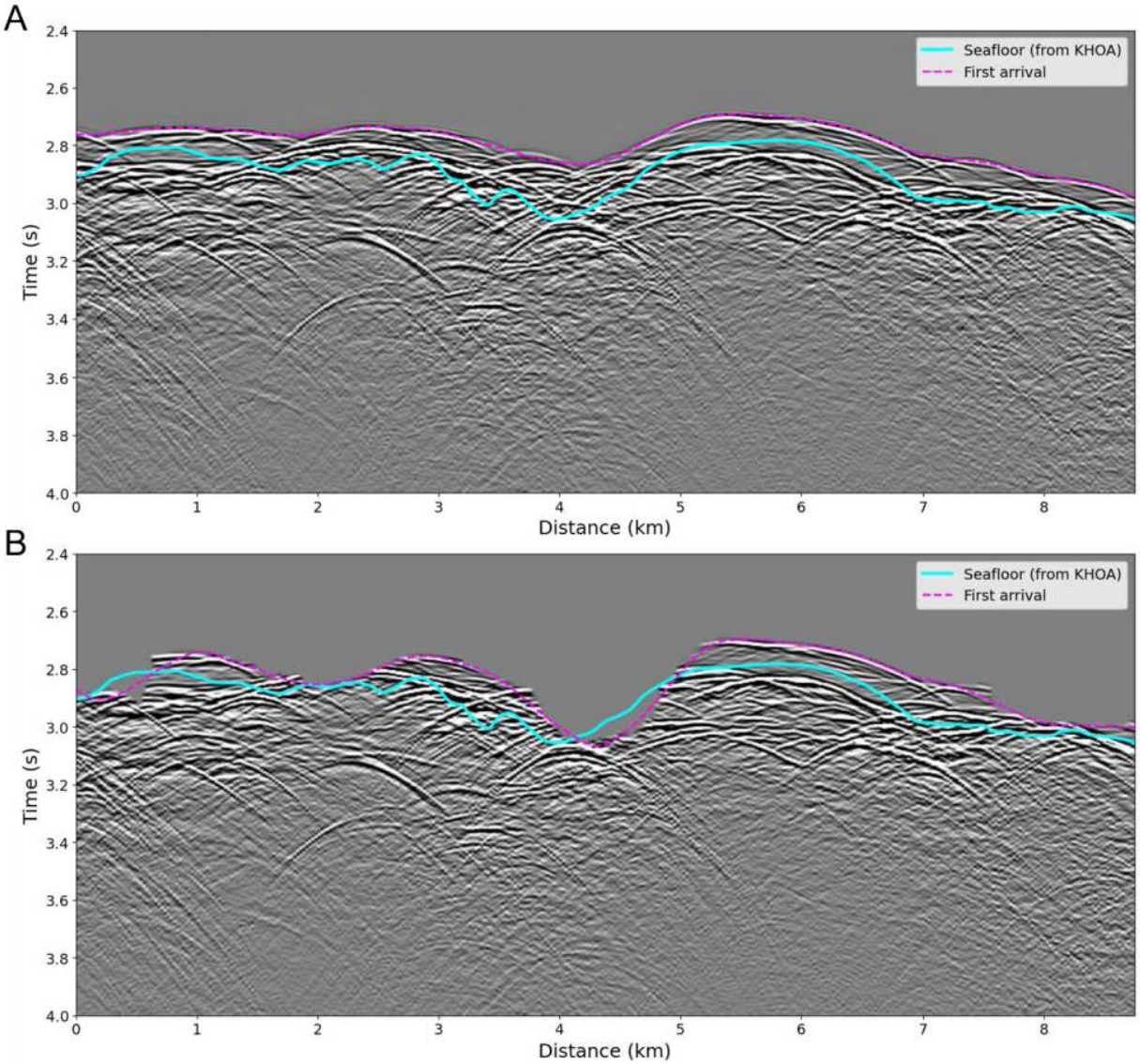



**Figure 16.** Definition of the travel time-delay metric with the external bathymetric reference from the Korea Hydrographic and Oceanographic Agency (KHOA). The KHOA profile (solid line) and the seafloor (dashed line) are overlaid on (A) the brute stack and (B) the header-muted stack.

subsurface image. Identifying and removing these signals is challenging, as they can be coherent and mimic true reflections. Conventional processing techniques, such as simple time gating or frequency-wavenumber filtering, often struggle to effectively isolate and remove OPS without damaging the desired primary reflections, particularly in 2D datasets where the out-of-plane origin cannot be resolved geometrically. As the F–K benchmark in this study demonstrates, the overlap of OPS and the seafloor reflection renders purely dip-based separation ineffective. Simple muting based on header information, like the water bottom time, was ineffective in this case because the recorded time was itself contaminated by the strong OPS arrivals.

To address this, we sought a method to accurately define a mute function that follows the true seafloor reflection, effectively removing the preceding OPS. We found that NN picking provided a suitable tool for this task. By training the NN on a small set of manually identified true seafloor arrivals within the affected zone, the network learned to recognize and pick this specific event across the dataset, even in the presence of the strong, interfering OPS.

**Table 5. Mean absolute error and standard deviation of the travel time difference between the seismic seafloor pick and an independent bathymetric reference.**

| Data | Mean absolute error (ms) | Standard deviation (ms) |
|---|---|---|
| Brute stack (Figure 4A) | 115.14 | 32.88 |
| Brute stack after removing out-of-plane signal (OPS) using header mute (Figure 4B) | 70.99 | 41.45 |
| F–K filtered stack section in common shot gather (CSG) (Figure 6A) | 120.40 | 37.28 |
| F–K filtered stack section in common receiver gather (CRG) (Figure 6B) | 134.95 | 41.20 |
| F–K filtered stack section in common depth point (CDP)–offset gather (Figure 6C) | 215.53 | 49.49 |
| Stack section after removing OPS in CSG (Figure 14A) | 49.29 | 33.41 |
| Stack section after removing OPS in CRG (Figure 14B) | 45.37 | 27.09 |
| Stack section after removing OPS in CDP–offset gather (Figure 14C) | 61.24 | 35.87 |
| Migrated section after removing OPS in CSG (Figure 15A) | 47.33 | 26.83 |
| Migrated section after removing OPS in CRG (Figure 15B) | 50.37 | 37.64 |
| Migrated section after removing OPS in CDP–offset gather (Figure 15C) | 48.31 | 30.86 |

Notes: Results are shown for the brute stack, header-muted stack, the neural network-muted stacks, and migrated sections from the three domains, and the F–K filtered stacks. Lower values indicate better agreement with the reference.

The NN picking approach offers several advantages for this type of targeted noise removal. It is adaptive, learning the specific characteristics of the target event directly from the data. It proved effective even with minimal training data, avoiding the need for large, pre-labeled datasets often required by more complex machine learning models. This study demonstrated the successful application of a technique commonly used for first-arrival picking in land seismic surveys to a different challenge in marine seismic data. While many automated picking algorithms focus solely on first breaks, the NN picker used here could be trained to target a specific later arrival (the true seafloor) based on user guidance. This adaptability makes it a flexible tool for addressing specific noise problems where a particular event needs to be identified and isolated.

While 3D seismic acquisition inherently handles OPS more effectively by resolving the wavefield in three dimensions, this is often not feasible due to cost constraints, especially when working with legacy 2D data. For such 2D datasets, particularly those acquired with an adequate number of channels (longer streamers provide better pattern definition), the NN picking method offers a practical and effective way to mitigate OPS. The results suggest that the quality of OPS removal improves with better data redundancy.

The ability to produce a clearer image of the true seafloor has significant practical implications. Accurate seafloor imaging is essential for various applications, including geohazard assessment, planning routes for submarine cables or pipelines, and selecting sites for ocean bottom instrumentation. By removing masking OPS, this technique can improve the reliability of interpretations

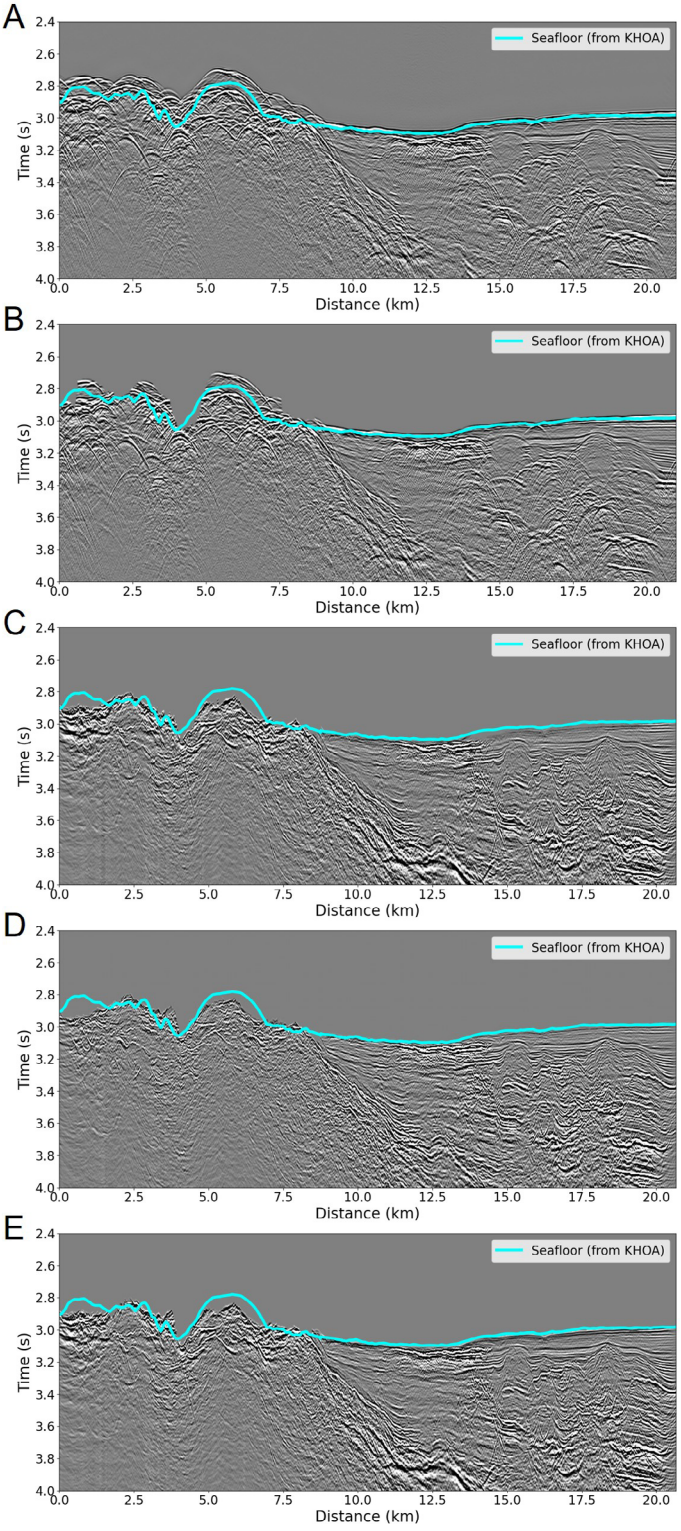



**Figure 17.** Validation by overlaying the Korea Hydrographic and Oceanographic Agency (KHOA) bathymetry profile (cyan line) on seismic sections. (A) The brute stack shows a poor match due to out-of-plane signals. (B) The header-muted stack shows some improvement, but the residual out-of-plane signal remains. (C–E) The final migrated sections after neural network-muting in the (C) common shot, (D) common receiver, and (E) common depth point–offset domains show a much closer alignment with the KHOA profile.

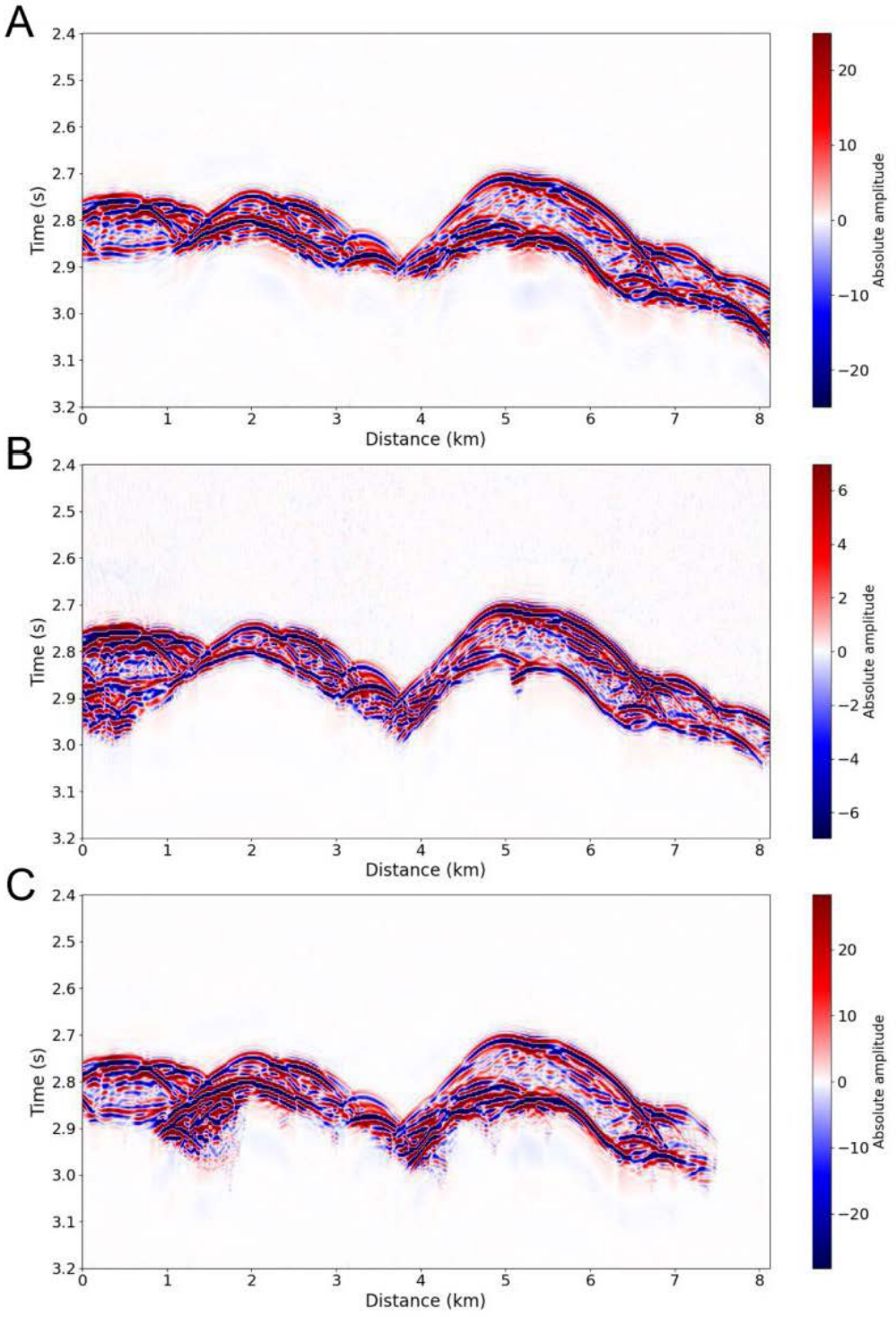



**Figure 18.** Different sections showing the energy removed by the neural network-derived mute. The panels show the absolute difference between the brute stack (Figure 4A) and the muted stacks from the (A) common shot, (B) common receiver, and (C) common depth point–offset domains. The display is limited to 2.4–3.2 s and 0–8 km to emphasize the out-of-plane signal energy targeted by the workflow. The coherent energy corresponds to the removed out-of-plane signal.

based on 2D seismic data for these purposes.

Despite its effectiveness, the proposed method has limitations that suggest directions for future work. Recent advances in deep learning have introduced strategies such as transfer learning (fine-tuning) and few-shot learning to reduce the dependence on large, fully labeled seismic datasets. For example, fine-tuning a pre-trained convolutional neural network has been shown to substantially reduce manual labeling requirements for first-break picking while maintaining competitive accuracy across real surveys.[36] However, multiple studies emphasize that model performance can be highly sensitive to out-of-distribution (domain-shift) conditions related to acquisition configuration and site-dependent noise characteristics, motivating explicit evaluation of cross-environment transferability.[37] Few-shot approaches have also been proposed to adapt seismic interpretation models to new surveys with minimal annotation, but they explicitly note that subsurface heterogeneity can limit label transferability across locations and require careful adaptation mechanisms.[38]

Despite these advances, applying transfer/few-shot paradigms to OPS suppression in legacy 2D marine data faces a critical domain-gap issue. OPS is often strongly site-specific coherent noise generated by unique local 3D bathymetric/topographic conditions (e.g., irregular seamounts), meaning that models pre-trained on generic datasets may experience negative transfer and fail to reliably separate OPS from true in-plane seafloor reflections. In this context, lightweight, interpreter-in-the-loop approaches that learn directly from a small number of site-specific examples can be preferable when the dominant challenge is a locally unique noise mode rather than a broadly recurring texture class.

The NN picking technique, while effective, is not without other limitations. Its success is highly dependent on the quality of the initial user picks, making it a semi-automated process requiring careful human oversight and quality control. This interpreter-in-the-loop design means the technique is not fully automated. An expert must decide on and pick the horizon on key gathers to guide the network. However, this involvement ensures the picked horizon remains geophysically plausible and consistent. Nevertheless, the workflow provides a valuable capability for removing noise that is difficult to address with conventional filtering, though it still requires user expertise. Careful application and validation are essential to ensure the uniqueness and exactness of the picked horizon and the resulting mute. Furthermore, where the OPS and true seafloor reflection are nearly coincident in time, full recovery of the primary reflection remains a challenge, as the muting process can result in localized over-muting or residual contamination, which contributes to the remaining discrepancies.

Although this study demonstrated the method on a single line, the workflow relies on a site-specific training strategy. This design allows the network to adapt to various geological environments and acquisition geometries by learning the unique morphological characteristics of the signal and noise in each new survey, provided a small set of representative training picks is supplied.

## 5. Conclusion

Rapid variations in seafloor morphology can generate significant OPS on 2D marine seismic data. Because OPS is not ubiquitous in all datasets, standardized removal workflows are less common compared to those for multiples or ghosts. To separate the OPS from the seafloor reflection signal, we focused on the first break picking function by NN picking. Then we transferred it to muting information of seismic traces. To compare mute performance, we trained three different gather domains. Among the three-gather domain muting, the common receiver gather domain result yielded the most consistent picks and the clearest final image of the true seafloor, and it had better results than conventional F–K filtering. Quantitative validation against the KHOA reference data (bathymetry) supports this conclusion. This demonstrates that OPS muting with NN picking, when applied thoughtfully in the appropriate data domain, can be a powerful and practical tool for improving the quality of legacy 2D seismic data suffering from out-of-plane signal contamination.

This methodology could be incorporated into a dedicated workflow for processing 2D seismic data acquired in areas known for irregular seafloor topography, such as near submarine volcanoes or mid-ocean ridges, where OPS is likely to be problematic. Future work could involve testing more advanced machine learning models for signal discrimination, particularly focusing on separating overlapping OPS and primary reflections. Furthermore, the principle of using NN picking to identify specific target horizons amidst noise could potentially be adapted for discriminating other challenging seismic events, such as identifying specific stratigraphic layers in challenging zones or improving first-break picking when early arrivals are contaminated by non-seismic noise sources. We found this application of OPS muting with NN picking, adapted from its typical use, to be a valuable technique improving seafloor imaging.

## Acknowledgments

None.

## Funding

This research was supported by the Basic Research Projects (GP2025-021 and GP2025-025) of the Korea Institute of Geoscience and Mineral Resources (KIGAM), funded by the Ministry of Science and ICT of the Republic of Korea.

## Conflict of interest

The authors declare that they have no competing interests.

## Author contributions

*Conceptualization:* Ganghoon Lee, Snons Cheong
*Formal analysis:* Ganghoon Lee, Snons Cheong
*Investigation:* Ganghoon Lee, Changyoon Lee, Junseok Kwon, Snons Cheong
*Methodology:* Snons Cheong
*Validation:* Ganghoon Lee, Changyoon Lee, Junseok Kwon, Snons Cheong
*Visualization:* Ganghoon Lee, Changyoon Lee, Snons Cheong
*Writing–original draft:* Ganghoon Lee, Changyoon Lee, Junseok Kwon, Snons Cheong
*Writing–review & editing:* Ganghoon Lee, Snons Cheong

## Availability of data

Data associated with this research requires permission from the Korea Institute of Geoscience and Mineral Resources (KIGAM) and can be obtained by contacting the corresponding author.

## References

1. Sheriff RE, Geldart LP. *Exploration Seismology*. Cambridge University Press; 1995.
2. Yilmaz Ö. *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Society of Exploration Geophysicists; 2001.

   doi: 10.1190/1.9781560801580
3. Drummond BJ, Hobbs RW, Goleby BR. The effects of out-of-plane seismic energy on reflections in crustal-scale 2D seismic sections. *Tectonophysics*. 2004;388(1-4):213-224.

   doi: 10.1016/j.tecto.2004.07.040
4. Larner K, Chambers R, Yang M, Lynn W, Wai W. Coherent noise in marine seismic data. *Geophysics*. 1983;48(7):854-886.

   doi: 10.1190/1.1441516
5. Hobbs RW, Drummond BJ, Goleby BR. The effects of three-dimensional structure on two-dimensional images of crustal seismic sections and on the interpretation of shear zone morphology. *Geophys J Int*. 2006;164(3):490-500.

   doi: 10.1111/j.1365-246x.2006.02814.x
6. Günther RH, Klemperer SL, Goodliffe AM. Modeling sideswipe in 2D oceanic seismic surveys from sonar data: Application to the Mariana arc. *Tectonophysics*. 2006;420(1-2):333-343.

   doi: 10.1016/j.tecto.2006.01.027
7. Tucker PM. *Pitfalls Revisited*. Society of Exploration Geophysicists; 1982.

   doi: 10.1190/1.9781560802341
8. Nedimović MR, West GF. Crooked-line 2D seismic reflection imaging in crystalline terrains: Part 1, data processing. *Geophysics*. 2003;68(1):274-285.

   doi: 10.1190/1.1543213
9. Nedimović MR, West GF. Crooked-line 2D seismic reflection imaging in crystalline terrains: Part 2, migration. *Geophysics*. 2003;68(1):286-296.

   doi: 10.1190/1.1543214
10. U.S. Energy Information Administration. Definitions, sources, and explanatory notes. Independent Statistics and Analysis. Available from: https://www.eia.gov/dnav/ng/TblDefs/ng_enr_seis_tbldef2.asp [Last accessed on 2025 Nov 20].
11. Biondi BL. *3D Seismic Imaging*. Society of Exploration Geophysicists; 2006.

    doi: 10.1190/1.9781560801689
12. Vermeer GJO. *3D Seismic Survey Design*. 2nd ed. Society of Exploration Geophysicists; 2012.

    doi: 10.1190/1.9781560803041
13. Micenko M. Seismic window: Pitfalls revisited. *Preview*. 2019;2019(200):40.

    doi: 10.1080/14432471.2019.1621248
14. Stewart RR, Marchisio G. Side-scanning seismic: Analysis and a physical modeling study. In: *SEG Technical Program Expanded Abstracts 1991*. 61st SEG Annual International Meeting; November 10-14, 1991; Houston, TX. Society of Exploration Geophysicists; 1991:679-682.

    doi: 10.1190/1.1889180
15. Duren RE, Morris SV. Sideswipe removal via null steering. *Geophysics*. 1992;57(12):1623-1632.

    doi: 10.1190/1.1443229
16. Beckel RA, Juhlin C. The cross-dip correction as a tool to improve imaging of crooked-line seismic data: a case study from the post-glacial Burträsk fault, Sweden. *Solid Earth*. 2019;10(2):581-598.

    doi: 10.5194/se-10-581-2019
17. Mancuso C, Naghizadeh M. Generalized cross-dip moveout correction of crooked 2D seismic reflection surveys. *Geophysics*. 2021;86(4):V285-V298.

doi: 10.1190/geo2020-0278.1

18. Alkhalifah TA, Biondi B. Numerical analysis of the azimuth moveout operator for vertically inhomogeneous media. *Geophysics*. 2004;69(2):554-561.

doi: 10.1190/1.1707075

19. Khoshnavaz MJ. Oriented time-domain dip moveout correction for planar reflectors in common-source domain. *Geophysics*. 2017;82(6):U87-U97.

doi: 10.1190/geo2016-0577.1

20. Korea Hydrographic and Oceanographic Agency. Bathymetric Data (BADA2024). Badanuri Marine Information Service. Available from: https://www.khoa.go.kr/oceandata/oceaninfo/map.do?lang=en [Last accessed on 2026 Mar 18].

21. Lee GH, Kim HJ, Han SJ, Kim DC. Seismic stratigraphy of the deep Ulleung Basin in the East Sea (Japan Sea) back-arc basin. *Mar Pet Geol*. 2001;18(5):615-634.

doi: 10.1016/s0264-8172(01)00016-2

22. Ryu BJ, Riedel M. Gas hydrates in the Ulleung Basin, East Sea of Korea. *Terr Atmos Ocean Sci*. 2017;28(6):943-963.

doi: 10.3319/tao.2017.10.21.01

23. Murat ME, Rudman AJ. Automated first arrival picking: A neural network approach. *Geophys Prospect*. 1992;40(6):587-604.

doi: 10.1111/j.1365-2478.1992.tb00543.x

24. McCormack MD, Zaucha DE, Dushek DW. First-break refraction event picking and seismic data trace editing using neural networks. *Geophysics*. 1993;58(1):67-78.

doi: 10.1190/1.1443352

25. Pu Y, Zhang X. Application of deep learning in first break picking of seismic data. In: *Proceedings of the SEG 2018 Workshop: SEG Maximizing Asset Value Through Artificial Intelligence and Machine Learning*. SEG 2018 Workshop; September 17-19, 2018; Beijing, China. Society of Exploration Geophysicists and Chinese Geophysical Society; 2018:19-21.

doi: 10.1190/aiml2018-05.1

26. Tsai KC, Hu W, Wu X, Chen J, Han Z. Automatic first arrival picking via deep learning with human interactive learning. *IEEE Trans Geosci Remote Sens*. 2020;58(2):1380-1391.

doi: 10.1109/tgrs.2019.2946118

27. Bornstein T, Lange D, Münchmeyer J, *et al*. PickBlue: Seismic phase picking for ocean bottom seismometers with deep learning. *Earth Space Sci*. 2023;11(1):e2023EA003332.

doi: 10.1029/2023ea003332

28. Niksejel A, Zhang M. OBSTransformer: a deep-learning seismic phase picker for OBS data using automated labelling and transfer learning. *Geophys J Int*. 2024;237(1):485-505.

doi: 10.1093/gji/ggae049

29. Pham N, Li W. Physics-constrained deep learning for ground roll attenuation. *Geophysics*. 2021;87(1):V15-V27.

doi: 10.1190/geo2020-0691.1

30. Cui Y, Wu J, Bai M, Chen Y. Ground-truth-free deep learning for 3D seismic denoising and reconstruction with channel attention mechanism. *Geophysics*. 2024;89(6):V503-V520.

doi: 10.1190/geo2023-0592.1

31. Fu X, Cao D, Zhang F. Deep-learning-based denoising of prestack seismic angle gathers with application to time-lapse data. *Geophysics*. 2025;91(1):WA11-WA21.

doi: 10.1190/geo2024-0472.1

32. Fahlman SE, Lebiere C. The cascade-correlation learning architecture. In: Touretzky D, ed. *Advances in Neural Information Processing Systems 2*. NIPS'89: Neural Information Processing Systems; November 27-30, 1989; Denver, CO. MIT Press; 1990:524-532.

33. Li Z, Sheng R, Xu W, Yang J. The improvement of neural network cascade-correlation algorithm and its application in picking seismic first break. *Adv Pet Explor Dev*. 2013;5(2):41-51.

doi: 10.3968/j.aped.1925543820130502.1503

34. Schneider WA. Integral formulation for migration in two and three dimensions. *Geophysics*. 1978;43(1):49-76.

doi: 10.1190/1.1440828

35. Bleistein N. On the imaging of reflectors in the Earth. *Geophysics*. 1987;52(7):931-942.

doi: 10.1190/1.1442363

36. Mardan A, Blouin M, Fabien-Ouellet G, Giroux B, Vergniault C, Gendreau J. A fine-tuning workflow for automatic first-break picking with deep learning. *Near Surf Geophys*. 2024;22(5):539-552.

doi: 10.1002/nsg.12316

37. St-Charles PL, Rousseau B, Ghosn J, Bellefleur G, Schetselaar E. A deep learning benchmark for first break detection from hardrock seismic reflection data. *Geophysics*. 2023;89(1):WA279-WA294.

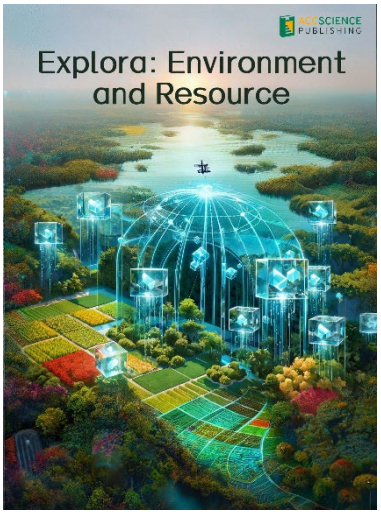
doi: 10.1190/geo2022-0741.1

38. Papadopoulos T, Mosser L, Oikonomou D, Naeini EZ, Karantzalos K. Few-shot Learning for Semantic Segmentation of Seismic Data. In: *82nd EAGE Annual Conference & Exhibition*. 82nd EAGE Annual Conference & Exhibition; October 18-21, 2021; Amsterdam, Netherlands. European Association of Geoscientists & Engineers; 2021:1-5.

doi: 10.3997/2214-4609.202113255

ACCSCIENCE PUBLISHING

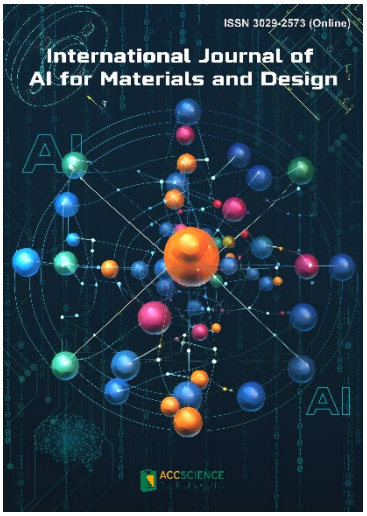
# OUR JOURNALS




*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
2 Venture Drive, #07-06 Vision Exchange, Singapore 608526
E-mail: editorial@accscience.com
Phone: +65 6348 3650