Petroleum Science 23 (2026) 7086–7103

Contents lists available at ScienceDirect

Petroleum Science

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

Original Paper

Seismic elastic parameter inversion via multi-attribute-assisted multi-trace supervised learning

Ren Luo, Huai-Zhen Chen, Jian-Hua Geng, Ben-Feng Wang *

State Key Laboratory of Marine Geology, Tongji University, Shanghai, 200092, China

ARTICLE INFO

Article history:
Received 31 December 2025
Received in revised form
29 April 2026
Accepted 19 May 2026
Available online 25 May 2026

Edited by Jie Hao

Keywords:
Prestack seismic inversion
Multi-attribute
Multi-trace
Elastic parameter
Supervised learning

ABSTRACT

Elastic parameters (primary wave velocity V_P, shear wave velocity V_S, and density) are crucial to link seismic data and the corresponding petrophysical properties. Traditional elastic parameter inversion methods require high quality seismic data and face challenges in the suitability of physical models, the decoupling of coupled parameters and high-resolution characterization. Deep learning method, as a nonlinear powerful tool, gains more attention for precise elastic parameter inversion. To improve the generalizability of the supervised optimization model, we propose a multiple attributes (partial angle stacks, partial angle stack differences and elastic impedance) assisted elastic parameter inversion method. The input data utilize amplitude variation characteristics from both seismic data and elastic impedance. In terms of network training, we first design an overlapping patching strategy during data preparation to augment the labeled data. Then, a special Unet, which performs max-pooling only along the time axis and employs a strip-like Conv2d kernel, is used for multi-trace inversion, during which a masking strategy is introduced to cope with single-trace labels. Besides, we employ attribute filtering and modified tanh activation function strategies to improve the performance in accuracy and horizontal continuity. Both synthetic and field data tests demonstrate that the proposed multi-attribute-assisted multi-trace elastic parameter inversion method exhibits superior accuracy, robust noise resistance, and satisfactory horizontal continuity.

© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Prestack seismic inversion can obtain multiple elastic parameters, such as primary wave velocity (V_P), shear wave velocity (V_S), and density. These parameters allow for a precise description of subsurface reservoir distribution. The seismic-derived elastic parameters are helpful to predict reservoir properties (Zou et al., 2023; Bosch et al., 2010; Han et al., 1986). Geophysicists have successfully utilized various combinations of elastic parameters to derive reservoir and fluid-related indicators, significantly enhancing oil and gas reservoir detection (Smith and Gidlow, 1987; Goodway et al., 1997). Therefore, high-precision prestack elastic parameter inversion is of paramount importance for reservoir and fluid prediction.

The Zoeppritz equations serve as a fundamental cornerstone for prestack elastic parameter inversion, and they elucidate the reflection and transmission characteristics of seismic waves at elastic interfaces. Considering the strongly nonlinear features of the Zoeppritz equations, Aki and Richards (1980) propose a linearized approximation based on the relative changes of elastic parameters. Although the Zoeppritz equations are proposed earlier, they are not properly used until the 1980s when Ostrander (1984) formally uses the characteristic of Amplitude Variation with Offset (AVO) to identify gas-bearing sandstone in field cases. Since then, a series of linear approximations of the Zoeppritz equations are proposed, such as Shuey equation (Shuey, 1985) and Fatti equation (Fatti et al., 1994). Connolly (1999) proposes the concept of elastic impedance (EI), which is analogous to acoustic impedance (AI), and derives the relationship between EI and V_P, V_S, and density from a linearization of the Zoeppritz equation. Based on the Fatti equation, a modification is proposed based on the relationships of elastic parameters from log data to enhance the inversion stability (Hampson et al., 2005). Given that the prestack angle gathers

* Corresponding author.
E-mail address: wbf1232007@tongji.edu.cn (B.-F. Wang).

Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.05.029

1995-8226/© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

exhibit a relatively low signal-to-noise ratio (SNR), researchers have significant interest in partial angle stacks (PAS) during elastic parameter inversion (Barens and Hansen, 2006). The PAS with improved SNR preserves AVO characteristics of prestack seismic data, and it can help improve the inversion performance and stability of elastic parameters (Wang, 2005). Currently, incorporating different constraints in the objective function, such as contributions of difference angle gathers (Wang et al., 2021b), cross-gradient constraints (Dai et al., 2022) and gradient structure similarity (Wang et al., 2023), has become a key research direction of AVO inversion. Another type of elastic parameter inversion method that uses elastic impedance (EI) as the intermediate variable is also widely applied (Mallick, 2001; Ma et al., 2011; Zong et al., 2012). To avoid cumulative errors caused by indirect calculation of reservoir-sensitive parameters from inverted elastic properties, researchers have increasingly focused on direct inversion frameworks. For example, Chen et al. (2022) derived a novel linearized PP-wave reflection coefficient equation directly linked to the P-to-S wave velocity ratio.

In recent years, deep learning (DL) methods have gained widespread usage in seismic inversion. Initially, supervised learning algorithms are used to establish a direct nonlinear relationship between the input seismic data and target parameters based on abundant labeled data (Cao et al., 2018; Zhong et al., 2019). Due to the impact of exploration costs, the quantity of labeled logging data is small, making it difficult for a pure seismic data-driven approach to solve complex inversion problems.

Nowadays, more and more researchers focus on prior constraints (Wu et al., 2023). In the area of inversion, physics-guided training strategies are widely studied. Given the challenges in acquiring labeled logging data, semi-supervised learning strategies are proposed to make full use of unlabeled data. Model-data dual-driven inversion is a prevalent semi-supervised strategy. Biswas et al. (2019) employ the Aki-Richards approximation as a forward model, and the mismatch between the observed data and the simulated data using DL-based elastic parameter estimation is regarded as another constraint during the network optimization. Sun et al. (2023) propose a model-data-driven generative adversarial network for prestack elastic parameter inversion. Similar approaches are widely utilized for estimating various geophysical parameters, including acoustic impedance (Wang et al., 2022) and anisotropy parameters (Li et al., 2025). Besides, replacing physical forward modeling with a forward network is another effective method. Wang et al. (2020) propose a semi-supervised closed-loop network framework for seismic impedance inversion. Its primary purpose is to harness the potential of unlabeled data, bolster the horizontal continuity of inversion results, and enhance the inversion accuracy (Wang et al., 2021a). Zhang et al. (2026) introduce a double dual neural network framework for prestack seismic inversion and integrate model-domain priors to enhance inversion robustness. However, semi-supervised learning algorithms also face certain limitations in practical applications. For instance, the used physical forward model might not accurately capture the nonlinear relationship between elastic parameters and seismic data. Replacing the physical forward model with a nonlinear network sacrifices the physical interpretability and may yield unstable or counterintuitive inversion results. Besides, the training efficiency is reduced because both labeled and unlabeled data are used for training.

To avoid errors caused by over-reliance on physical forward modeling in semi-supervised inversion while ensuring the model's generalization performance, it is essential to prioritize multiple input attributes. Incorporating the initial model as part of the training input can improve the horizontal continuity and inversion accuracy of the estimated results (Wu et al., 2021; Song et al., 2022; Zhang et al., 2023). Furthermore, researchers employ multiple seismic attributes assisted algorithms for lithology prediction (Dunham, 2021; Xu et al., 2022; Zhao et al., 2026). These algorithms utilize angle gathers and seismic attributes as the network input to enhance the lithology prediction accuracy. Currently, the majority of multi-attribute DL applications in geophysics focus on classification tasks, such as facies classification and lithology prediction, whereas research on multi-attribute-assisted seismic parameter inversion remains limited.

Our previous work included a preliminary study on multi-attribute inversion (Luo et al., 2023). In this paper, we propose a DL-based multi-attribute-assisted multi-trace algorithm for V_P, V_S, and density estimation in a supervised learning manner. Unlike multi-attribute lithofacies prediction approaches that rely on a large number of attributes, our method selects only three attributes that are theoretically closely related to elastic parameters. Except for PAS, the proposed method further utilizes partial angle stack differences (PASD) and the estimated EI attribute as the extra inputs of the network. Furthermore, we introduce a multi-trace strategy to enhance the inversion performance. Synthetic and field data examples demonstrate that our proposed method exhibits robust noise resistance, satisfactory horizontal continuity and superior accuracy.

2. Methodology

2.1. Review of traditional prestack inversion method

2.1.1. AVO inversion

The Aki-Richards approximation is widely used for elastic parameter estimation, and it is re-expressed in Eq. (1) by Fatti et al. (1994):

$$R(\theta) = c_1 R_P + c_2 R_S + c_3 R_D, \tag{1}$$

where $R(\theta)$ is angle dependent reflectivity, θ is the angle of incidence, $c_1 = 1 + \tan^2\theta$, $c_2 = -8\gamma^2 \sin^2\theta$, $c_3 = -0.5\tan^2\theta + 2\gamma^2 \sin^2\theta$, $\gamma = V_S/V_P$, and the reflectivity terms (R_P, R_S, R_D) are given by:

$$\begin{aligned} R_P &= \frac{1}{2}\left[\frac{\Delta V_P}{V_P} + \frac{\Delta\rho}{\rho}\right], \\ R_S &= \frac{1}{2}\left[\frac{\Delta V_S}{V_S} + \frac{\Delta\rho}{\rho}\right], \\ R_D &= \frac{\Delta\rho}{\rho}. \end{aligned} \tag{2}$$

Equation (2) can be approximated using a log-differential form and a seismic trace $S(\theta)$ can be simulated by convolving a seismic wavelet with reflectivity, given by:

$$S(\theta) = \frac{1}{2} W(\theta) D (c_1 L_P + c_2 L_S + 2c_3 L_D), \tag{3}$$

where $L_P = \ln(Z_P) = \ln(\rho V_P)$, $L_S = \ln(Z_S) = \ln(\rho V_S)$, $L_D = \ln(\rho)$, $W(\theta)$ is an angle dependent wavelet matrix and D is a derivative matrix given by:

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots \\ 0 & -1 & 1 & \ddots \\ 0 & 0 & -1 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{4}$$

Hampson et al. (2005) note that the relationship between L_P and L_S or L_D can be obtained roughly using a linear fitting method from well-logs, given by:

$$L_S = \alpha L_P + \alpha_c + \Delta L_S, \tag{5}$$

and

$$L_D = \beta L_P + \beta_c + \Delta L_D, \tag{6}$$

where α and β are slopes, α_c and β_c are intercepts, and the terms of ΔL_S and ΔL_D are the deviations away from the linear fitting that we aim to solve for.

Combining Eqs. (3)–(6), the expression of $S(\theta)$ can be modified as:

$$S(\theta) = W(\theta)D(\tilde{c}_1 L_P + \tilde{c}_2 \Delta L_S + c_3 \Delta L_D), \tag{7}$$

where $\tilde{c}_1 = (c_1 + \alpha c_2)/2 + \beta c_3$ and $\tilde{c}_2 = c_2/2$. Equation (7) utilizes the prior information of subsurface parameters at the well locations for V_P, V_S, and density estimation. Besides, the ranges of ΔL_S and ΔL_D can be set as constraints to ensure inversion stability. The inversion results based on Eq. (7) can be obtained using Hampson Russell software.

2.1.2. EI inversion

The EI is first given by Connolly (1999):

$$\mathrm{EI}(\theta) = V_P^{(1+\tan^2\theta)} V_S^{-8\gamma^2 \sin^2\theta} \rho^{(1-4\gamma^2 \sin^2\theta)}. \tag{8}$$

Whitcombe (2002) proposes a normalized EI to avoid its rapid variation with incident angle:

$$\mathrm{EI}(\theta) = V_{P0}\rho_0 \left[\left(\frac{V_P}{V_{P0}}\right)^{(1+\tan^2\theta)} \left(\frac{V_S}{V_{S0}}\right)^{-8\gamma^2 \sin^2\theta} \left(\frac{\rho}{\rho_0}\right)^{(1-4\gamma^2 \sin^2\theta)} \right], \tag{9}$$

where V_{P0}, V_{S0} and ρ_0 are reference constants.

Similar to AI estimation, EI can be estimated from PAS using inversion algorithms. By taking the natural logarithm of both sides of Eq. (8) or (9), we can obtain a linear relationship between logarithmic EI and logarithmic elastic parameters. It allows us to perform least-square inversion for V_P, V_S, and density using the estimated EI (Ma et al., 2011; Zong et al., 2012). Compared with conventional AVO inversion, EI inversion has stronger anti-noise capability and robustness. However, it also faces challenges in multiple propagations of uncertainty.

The AVO and EI inversions for elastic parameters are well-established, grounded in clear physics, and highly efficient. However, traditional inversion methods face challenges in constructing satisfactory physical relationships between seismic data and elastic parameters. Besides, the resolution is dependent on bandwidth limited seismic data which always lack high-frequency information. DL algorithms can nonlinearly capture high-level features of seismic data for accurate characterization, thereby helping improve the inversion performance.

2.2. Multi-attribute assisted elastic parameter inversion

DL methods do not rely on classical physical models. Instead, they establish nonlinear mapping relationships between seismic data and target parameters in a data-driven manner, utilizing large volumes of training labels. However, well-logging data are expensive to achieve, and it is challenging to obtain large volumes of labels in field cases. Multiple attributes of seismic data can potentially improve the DL performance in the cases with limited labels. In this paper, we mainly concentrate on PAS, PASD and EI.

2.2.1. Multi-attribute extraction

Traditional DL-based methods mostly utilize prestack angle gathers (AG) as the network input for elastic parameter inversion. However, noise appearance in AG can significantly impair the inversion accuracy. To suppress noise effects and fully utilize the AVO characteristics and spatial attributes in the model domain, we design the following multi-attribute extraction procedure.

(1) Stack small, middle and large angles of prestack AG to obtain PAS $S(\theta_i)$ $(i = 1, 2, 3)$;
(2) Calculate the differences between PAS to obtain three PASD, which are defined as $\Delta S_1 = S(\theta_2) - S(\theta_1)$, $\Delta S_2 = S(\theta_3) - S(\theta_2)$ and $\Delta S_3 = S(\theta_3) - S(\theta_1)$;
(3) Estimate the EI attributes of different PAS using a model-based method (Russell, 1988) in Hampson Russell software to achieve $\mathrm{EI}(\theta_i)$;
(4) Carry out low-pass filtering (LPF) for ΔS_i and $\mathrm{EI}(\theta_i)$ to extract macro trend for a proper constraint.

We use PAS and the low-pass filtered PASD and EI as multiple attributes to improve the inversion performance. The specific reasoning can be illustrated as follows. Compared with the original prestack AG, PAS exhibits higher SNR and preserves the AVO characteristic. PASD is an advanced AVO attribute, which can be regarded as an approximation of AVO gradient. As EI is a model domain feature, it has the potential to improve inversion performance. Many studies have demonstrated that adding model domain features, such as initial models (Wu et al., 2021) and relative geologic time (Di et al., 2022), can improve the inversion performance. Besides, EI incorporates AVO anomaly information, which facilitates elastic parameter inversion. Another advantage is that we can obtain stable and physically interpretable EI using traditional inversion methods to constrain the DL outcomes. The obtained results after LPF have low frequency macro characteristics, enhancing the anti-noise capability and achieving more stable inversion results when using DL algorithms. For better illustrations, Fig. 1 shows the workflow of the multi-attribute extraction procedure. As the training input is ready, we can design inversion networks for elastic parameter estimation.

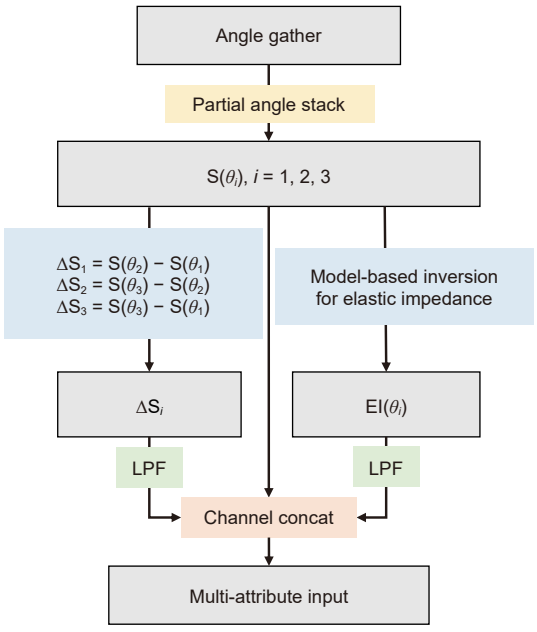


Fig. 1. The designed multi-attribute extraction workflow.

2.2.2. The single-trace inversion network

To mitigate the impact of numerical ranges, Z-score normalization (Shalabi et al., 2006) is applied to the inputs (PAS, PASD, EI), and the initial models and labels are scaled by dividing by their corresponding maximum values. The input for network training has a size of (N_l, N_t, N_n), where N_l, N_t, N_n denote the number of labels, the number of time samples and the number of channels. For a specific Common Depth Point (CDP), three PAS (small, middle, large angles stacks), and the corresponding difference attributes and EI attributes after LPF are combined along the channel, thus $N_n = 9$. We define the network without considering the batch dimension. The convolutional layer symbol is defined as $F_{C_1 \to C_2}^{k_1 \times k_2}$, where C_1 is the number of input channels and C_2 is the number of output channels, k_1 and k_2 denote the kernel sizes. When $k_2 = 1$, it represents 1D convolution. For our single-trace network, the dimensions of the input and the output of the convolution layer are $N_t \times C_1$ and $N_t \times C_2$, respectively. The first layer can be expressed as:

$$\mathbf{x}_2 = \mathrm{ReLU}\left(F_{9\to16}^{3\times1}(\mathbf{x}_1)\right), \tag{10}$$

where $\mathrm{ReLU}(x) = \max(0, x)$. Then there are 5 residual blocks, which are given by:

$$\mathbf{x}_{i+1} = \mathrm{Res}(\mathbf{x}_i) = \mathrm{ReLU}\left(F_{16\to16}^{3\times1}\left(\mathrm{ReLU}\left(F_{16\to16}^{30\times1}(\mathbf{x}_i)\right)\right) + \mathbf{x}_i\right), \tag{11}$$

where $i = 2, 3, \ldots, 6$. As V_P, V_S, and density have certain structural similarities, a multi-task learning strategy is used with mutual constraints. Specifically, some information of V_P can contribute to the estimation of V_S, and the information of V_P and V_S can contribute to density estimation. The network output is relative values of V_P, V_S, and density given by:

$$\begin{aligned}
\Delta V_P &= a\,\mathrm{Tanh}\left(F_{16\to1}^{3\times1}(\mathbf{x}_5)\right),\\
\Delta V_S &= a\,\mathrm{Tanh}\left(F_{16\to1}^{3\times1}(\mathbf{x}_6)\right),\\
\Delta\rho &= a\,\mathrm{Tanh}\left(F_{16\to1}^{3\times1}(\mathbf{x}_7)\right).
\end{aligned} \tag{12}$$

where $a\,\mathrm{Tanh}$ is a modified Tanh activation function that is illustrated in the following part.

After that, the relative values are fused with the initial models to achieve high-precision elastic parameter prediction. Fig. 2 shows the designed workflow diagram of the single-trace inversion network in detail. From our experiments, we found that the designed network structure outperforms the U-Net that uses single-trace input and Conv1d kernel. The network input consists of multi-attribute information extracted from a single CDP gather and the algorithm is called multi-attribute-assisted single-trace inversion (MSI). The loss function for MSI can be expressed as:

$$L_{\mathrm{MSI}} = \frac{1}{3N_t}\sum_{k=1}^{3}\sum_{t=1}^{N_t}\|y[t,k] - y^p[t,k]\|_2^2, \tag{13}$$

where $y[t,k]$ is a target elastic parameter at the time sample t, and $y^p[t,k]$ is the corresponding predicted value. When the parameter k equals 1, 2 or 3, it denotes V_P, V_S, and density, respectively. The network is optimized using the Adam optimizer with a warmup learning rate strategy (Goyal et al., 2018).

2.2.3. Modified tanh activation function

The network output represents relative values which often show large deviations when compared with the absolute values. Thus, a modified tanh activation function $a\,\mathrm{Tanh}(x)$ is designed to limit the output values to a reasonable range, as shown in Fig. 3. The modified activation function is employed at the end of the network which can enhance convergence speed and improve

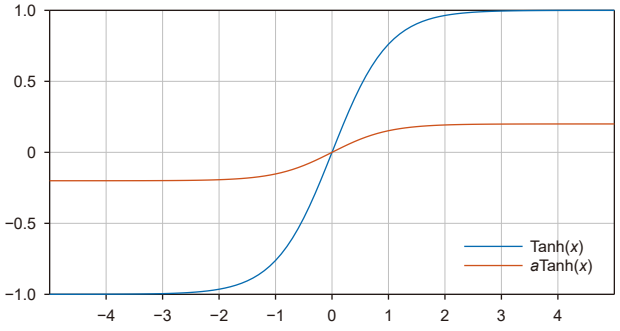


Fig. 3. Modified tanh activation function.

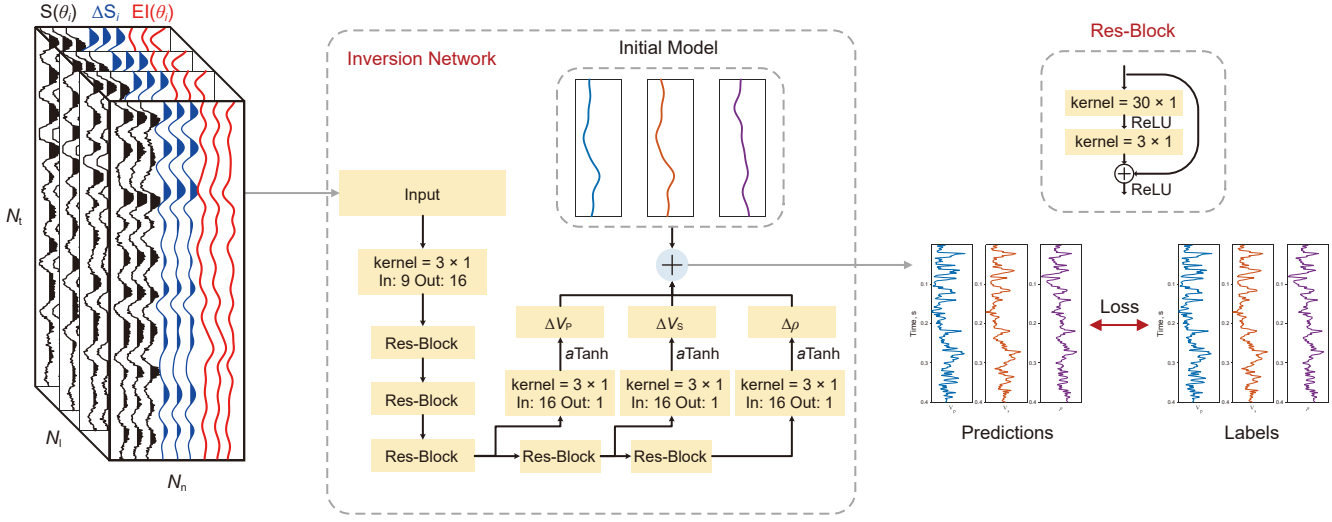


Fig. 2. The designed multi-attribute-assisted single-trace elastic parameter inversion workflow.

prediction accuracy. The minimum value of a can be estimated from normalized well labels using Eq. (14):

$$a_{\min} = \max\left(\left|\Delta m_t^i\right|\right), i = 1, 2, ..., N_l, t = 1, 2, ..., N_t, \tag{14}$$

where the symbol Δm_t^i denotes the relative value of V_P, V_S, and density for the tth time sample of the ith label. The value of a should be slightly larger than $a_{\min}$ to avoid the residual increase of peak values caused by the excessive constraint. The parameter $a_{\min}$ means the maximum allowable deviation, expressed as the maximum relative value of normalized V_P, V_S, and density.

2.2.4. Multi-trace elastic parameter inversion

The aforementioned MSI algorithm only uses one-dimensional convolution in the network architecture to capture the nonlinear relationship between the multi-attribute input and the desired output. It fails to fully exploit the horizontal continuity of extracted seismic attributes. Therefore, a multi-attribute-assisted multi-trace inversion (MMI) method is proposed, with a designed network framework shown in Fig. 4. The input is the multi-attribute information, while the output is the predicted relative values of elastic parameters, both including adjacent multiple CDPs.

After the multi-attribute extraction, we can obtain the input data with a size of $(N_t, N_x, 9)$. Here, the symbol N_x denotes the number of CDPs. Then a patching strategy is used to divide the original data into N_{TP} input patches $\mathbf{X}^p \in \mathbb{R}^{H\times W\times 9}(p = 1, 2, ...N_{TP})$, where H and W are the lengths of sliding windows in the time and CDP directions, respectively. Here, we set $H = N_t$ to avoid dividing in the time direction. When the sliding step is υ, the interval of the sliding window in the CDP direction is $[k\upsilon + 1, k\upsilon + W]$ with $k \in \mathbb{N}_0$. When $k\upsilon + W$ is greater than N_x, the interval of the window needs to be adjusted to $[N_x - W + 1, N_x]$. It should be mentioned that we avoid an excessively large W in the spatial direction, because we aim to accurately convey the high-frequency information of the well labels to the non-well locations through the spatial convolution operator. Consequently, we choose $W = 4$ in our numerical examples. Similarly, we divide the target data into patches $\mathbf{Y}^p \in \mathbb{R}^{H\times W\times 3}$. Given that the well label is single-trace while the output remains multi-trace, we should select the training data as the following steps:

(1) Construct zero-tensor with size $\mathbf{T} \in \mathbb{R}^{N_t\times N_x\times 3}$;
(2) Assign the value of one to the elements of T corresponding to the well-logging positions;
(3) Divide the tensor T into N_{TP} patches $\mathbf{M}^p \in \mathbb{R}^{H\times W\times 3}(p = 1, 2, ..., N_{TP})$ using the same patch-division method as for the input patches.
(4) Let $s^p = \sum_{k=1}^{3}\sum_{j=1}^{W}\sum_{i=1}^{H}\mathbf{M}_{i,j,k}^p$, which is the sum of all elements in the M^p patch. And find a set $\mathbf{P}_T = \{p|s^p > 0, p = 1, 2, ...N_{TP}\}$.
(5) Collect all pairs of input-target patches $\left(\mathbf{X}^t, \mathbf{Y}^t\right)$ for $t \in \mathbf{P}_T$ as training set and the corresponding M^t as the mask tensor.

The designed network adopts a U-Net architecture, which contains a series of convolution blocks, max-pooling layers, and up-convolution layers, as shown in Fig. 4. The convolution block can be expressed as:

$$\mathrm{ConB}_{C_{\mathrm{In}}}^{C_{\mathrm{Out}}}(\mathbf{X}) = \mathrm{ReLU}\left(F_{C_{\mathrm{Out}}\to C_{\mathrm{Out}}}^{K\times 3}\left(\mathrm{ReLU}\left(F_{C_{\mathrm{In}}\to C_{\mathrm{Out}}}^{K\times 3}(\mathbf{X})\right)\right)\right), \tag{15}$$

where C_{In} and C_{Out} are the input channels and output channels, respectively. Considering the shape of patches, we select a strip-like Conv2d kernel with a size of $(K, 3)$. The number of channels after the first Conv2d operation is set to C. The Discussion part provides the optimization strategy for parameters K and C. Besides, we only perform max-pooling and up-convolution in the time direction, considering that the parameter W is much smaller than

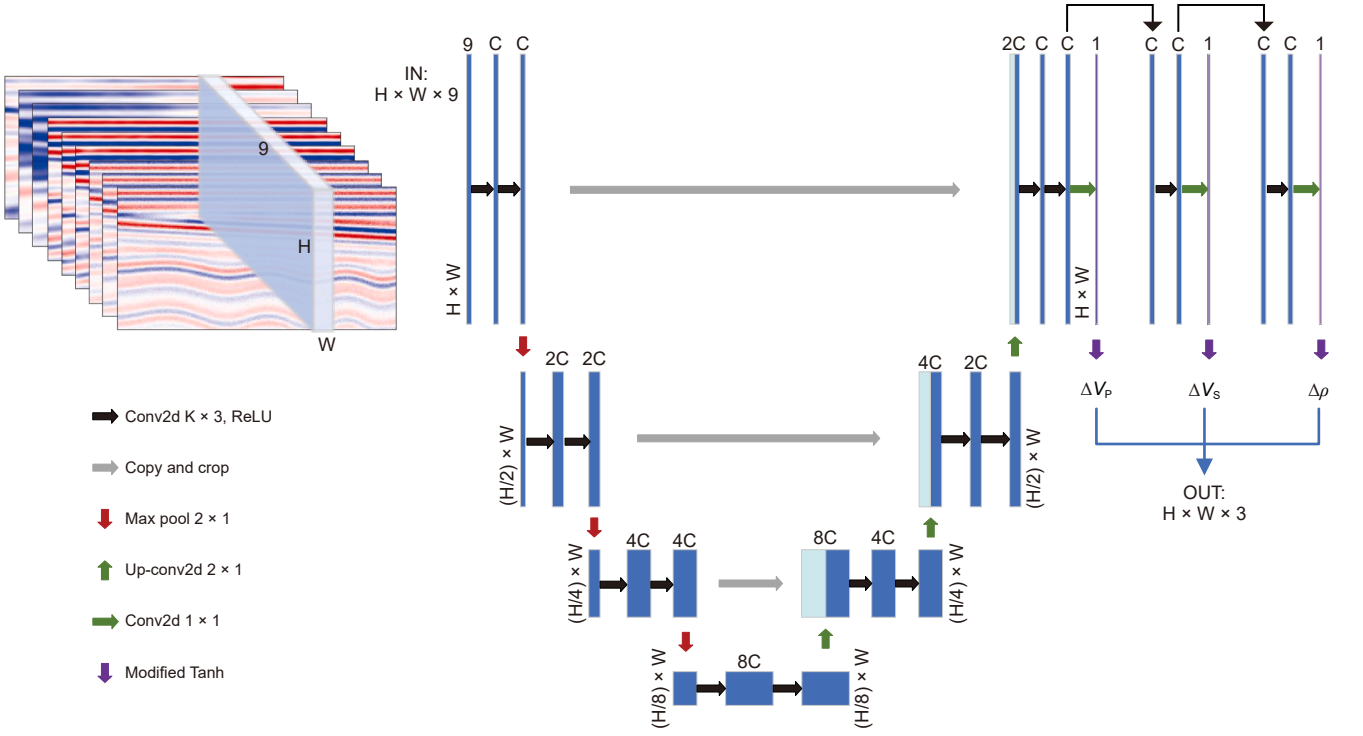


Fig. 4. The designed 2D U-Net with multi-attributes of adjacent multiple CDPs as the input and the corresponding elastic parameters of multiple CDPs as the output.

H. More details of the network structure can be obtained from Fig. 4. Similar to the strategy of MSI, the network outputs the relative values of three elastic parameters, given by:

$$\begin{aligned}\Delta V_{\mathrm{P}} &= a\,\mathrm{Tanh}\left(F_{C\to 1}^{1\times 1}(\mathbf{X}_{l1})\right),\\ \Delta V_{\mathrm{S}} &= a\,\mathrm{Tanh}\left(F_{C\to 1}^{1\times 1}(\mathbf{X}_{l2})\right),\\ \Delta\rho &= a\,\mathrm{Tanh}\left(F_{C\to 1}^{1\times 1}(\mathbf{X}_{l3})\right),\end{aligned} \tag{16}$$

where $\mathbf{X}_{l2} = \mathrm{ConB}_C^C(\mathbf{X}_{l1})$ and $\mathbf{X}_{l3} = \mathrm{ConB}_C^C(\mathbf{X}_{l2})$. The relative values are then added to the initial models to obtain the final estimation result $\mathbf{Y}_{\mathrm{est}}$.

The loss function for MMI can be expressed as:

$$L_{\mathrm{MMI}} = \frac{1}{\sum_{k=1}^{3}\sum_{j=1}^{W}\sum_{i=1}^{H}\mathbf{M}_{i,j,k}}\sum_{k=1}^{3}\sum_{j=1}^{W} \times \sum_{i=1}^{H}\left\|\mathbf{Y}_{i,j,k} - \mathbf{M}_{i,j,k}\odot(\mathbf{Y}_{\mathrm{est}})_{i,j,k}\right\|_2^2, \tag{17}$$

where $(\mathbf{Y}_{\mathrm{est}})_{i,j,k}$ and $\mathbf{Y}_{i,j,k}$ are the elements of estimated results and labels. The value of $\mathbf{Y}_{i,j,k}$ is equal to 0 at non-well locations. After the network training, all patches are input into the trained network to obtain the predictions $\mathbf{Y}_{\mathrm{est}}^p \in \mathbb{R}^{H\times W\times 3}(p = 1,2,\ldots,N_{\mathrm{TP}})$. We define a mapping relationship that maps the position (i, j) in the p-th patch to the corresponding position (t, x) in the original data space. For each patch index p from 1 to N_{TP}, and for each position (i, j) in the (H, W) grid of the p-th patch and each channel k from 1 to 3, we can update the integrated result tensor $\mathbf{Y}_{\mathrm{integrated}}$ and a temporary tensor C as:

$$\begin{aligned}\mathbf{Y}_{\mathrm{integrated}}(t,x,k) &= \mathbf{Y}_{\mathrm{integrated}}(t,x,k) + \mathbf{Y}_{\mathrm{est}}^p(i,j,k),\\ \mathbf{C}(t,x,k) &= \mathbf{C}(t,x,k) + 1,\end{aligned} \tag{18}$$

After integration processing, an overlapping estimated result can be obtained, and we need to average the overlapping parts to get the final elastic prediction results as:

$$\mathbf{Y}_{\mathrm{final}}(t,x,k) = \frac{\mathbf{Y}_{\mathrm{integrated}}(t,x,k)}{\mathbf{C}(t,x,k)}. \tag{19}$$

3. Numerical examples

In this section, we apply the MSI and MMI methods to synthetic data and field data. Some comparative methods (traditional method, AG-based method, and PAS-based method) are declared in advance. The results of the traditional method are obtained using Hampson Russell software (Hampson et al., 2005). The AG-based method and the PAS-based method utilize AG and PAS as inputs, respectively, and both of them are single-trace methods.

3.1. Synthetic data example

Fig. 5(a) shows 2D elastic parameter models for V_{P}, V_{S}, and density and Fig. 6(a) shows V_{P}, V_{S}, and density curves corresponding to CDP = 150. Initial models of V_{P}, V_{S}, and density are shown in Fig. 5(b), which are established via well-log interpolation along horizons. The labeled well-logs are located at the 50th, 150th, 250th, and 350th traces, as marked by the vertical lines in Fig. 5. Based on the Zoeppritz equations, a Ricker wavelet with a dominant frequency of 20 Hz is used to simulate AG via a convolution operator, as shown in Fig. 6(b). The size of AG is 401, 20, and 400 for the number of CDPs, the number of angles per CDP, and the number of time samples per angle, respectively. The angle increment is 2° from 2° to 40°. Random noise is added to the prestack AG, and the noisy gather with a SNR of 5 dB for CDP = 150 is shown in Fig. 6(c).

To reduce the influence of random noise, a partial stack operation is employed on the noisy prestack AG. The angles ranging from 2° to 14°, from 16° to 28° and from 30° to 40° are stacked to generate PAS. The central angles of the three PAS are 8°, 22° and 35°, respectively. Fig. 7 shows the extracted multi-attributes after Z-Score normalization.

As mentioned above, the size of MSI is (4, 400, 9). As for the MMI input, we set the sliding step υ to 1 in the spatial direction to generate 16 (4×4) labeled synthetic data, and the size of MMI is (16, 400, 4, 9). To highlight the advantages of the proposed MMI method, Fig. 8 shows the estimated elastic parameters using the traditional method, AG-based method, MSI method and MMI method for 5 dB noisy data. For the AG-based method, MSI method and MMI method, the modified tanh activation function is used for constraint. The calculated value of $a_{\min} = 0.1722$, thus $a = 0.2$ is an appropriate value. Fig. 8 shows that multi-attribute-assisted

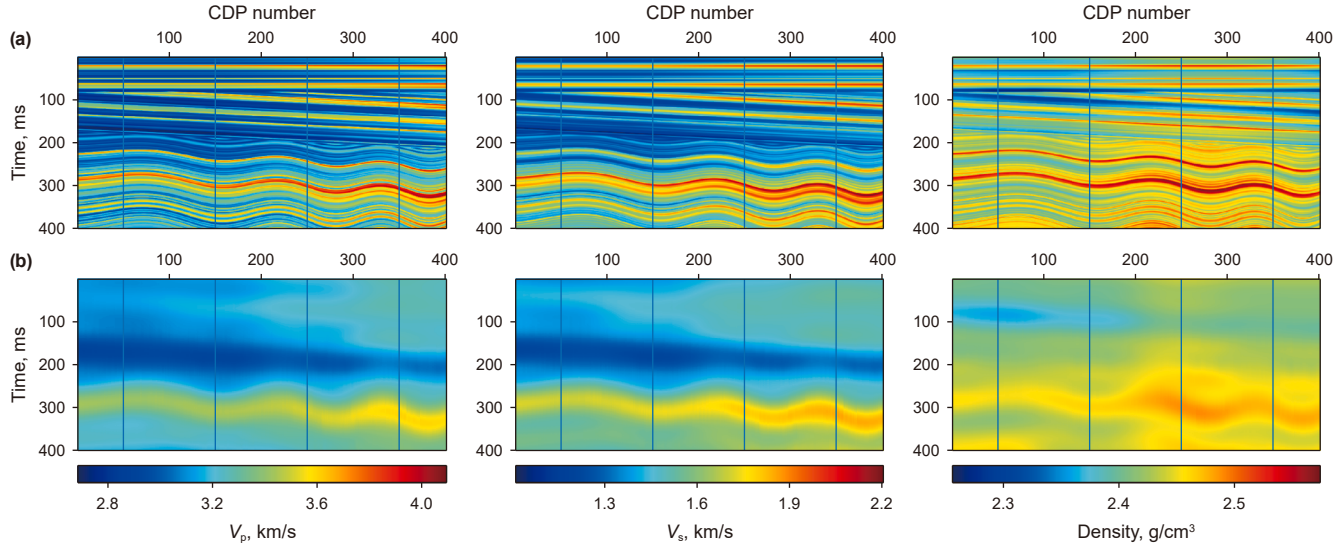


Fig. 5. Elastic parameter models. **(a)** and **(b)** are 2-D elastic parameter models of V_{P}, V_{S}, and density and the corresponding initial models.

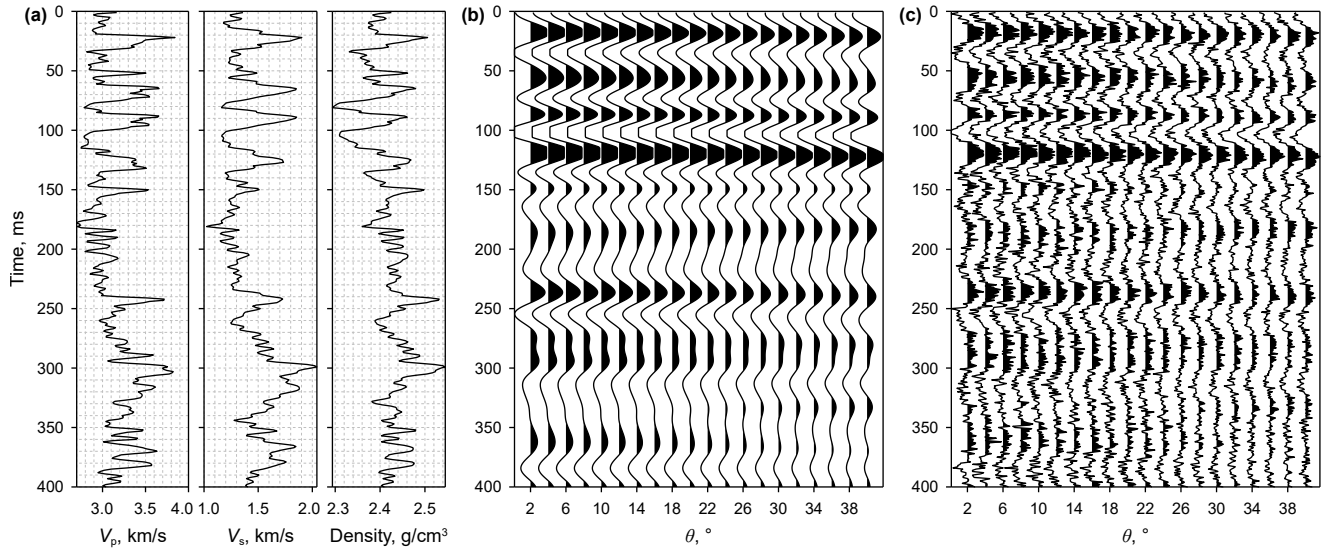


Fig. 6. The 1-D elastic parameter models and synthetic AG corresponding to CDP = 150. **(a)** V_P, V_S, and density curves; **(b)** and **(c)** are the synthetic AG without and with random noise (SNR = 5 dB).

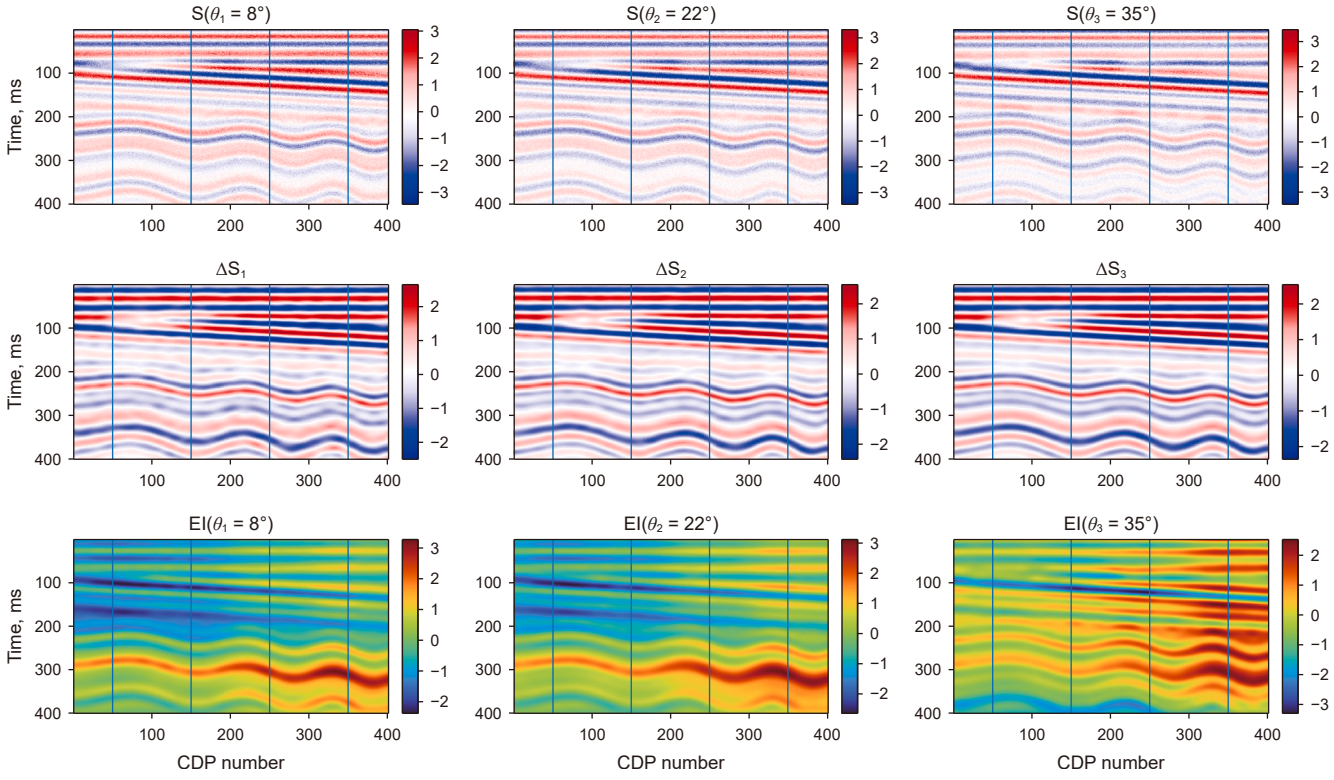


Fig. 7. The extracted multi-attributes after Z-Score normalization.

algorithms provide more accurate results. Fig. 9 shows the residuals between the estimated results and the true values, which further validates the effectiveness of the proposed MMI method with the least residual. The inversion results indicate that the traditional method has limited resolution and struggles to accurately characterize thin layers. The DL-based method can enhance resolution based on the contribution of available wells. However, the three DL-based methods mentioned exhibit different performance levels in terms of noise and lateral continuity. The AG-based method is highly sensitive to noise, leading to speckle noise and limited lateral continuity. The MSI method significantly suppresses the influence of noise, but residual speckle noise remains due to the limitation of single-trace algorithm. The MMI method minimizes the influence of noise, resulting in superior lateral continuity.

Table 1 quantitatively evaluates the estimated elastic parameters shown in Fig. 8 by utilizing the root mean square error (RMSE) metric. The results reveal a consistent trend: the RMSE values of

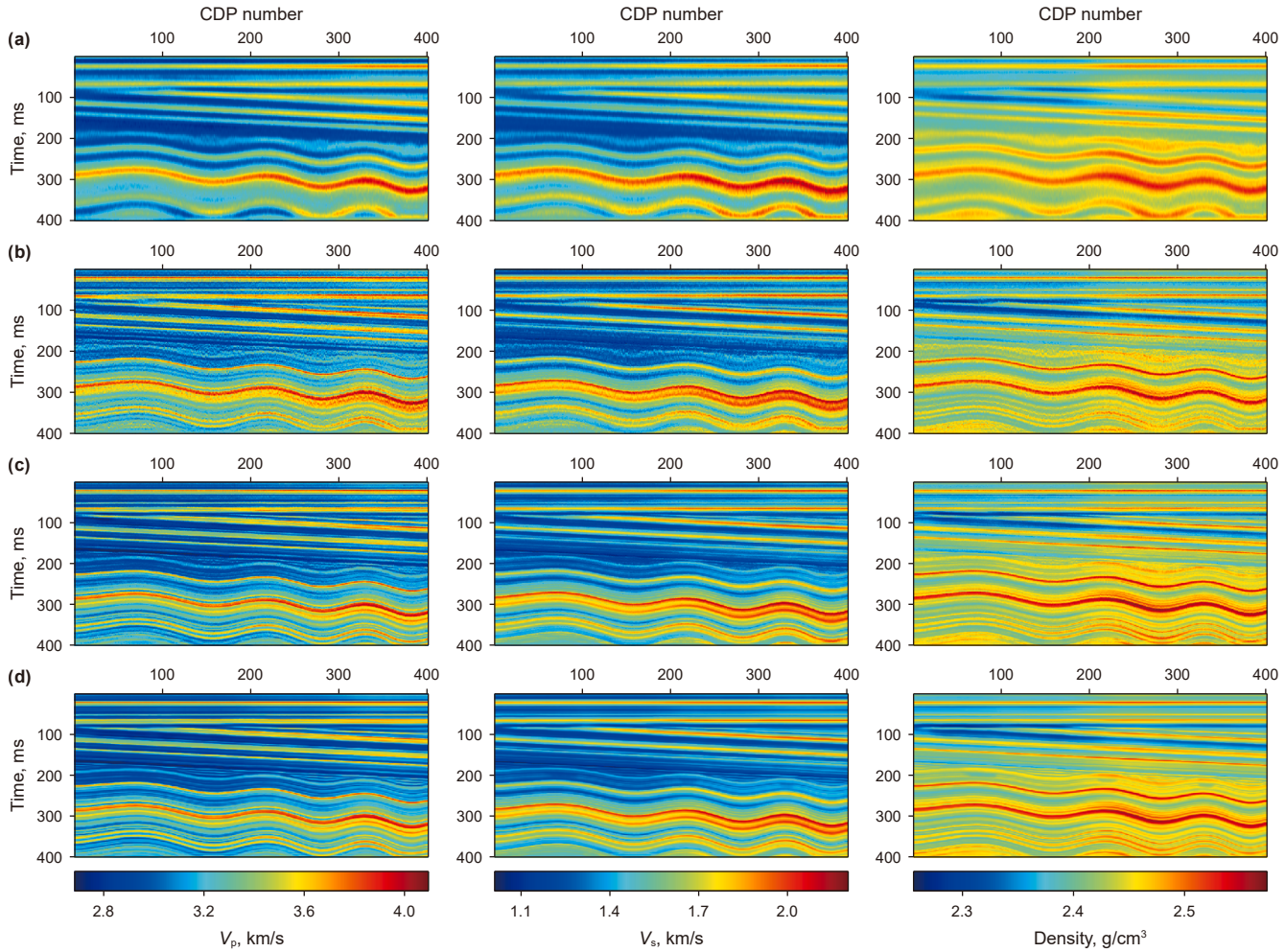


Fig. 8. Comparison of the estimated results. **(a)** to **(d)** represent the estimated results using the traditional method, AG-based method, MSI method and MMI method, respectively. The first, second and third columns represent V_P, V_S, and density, respectively.

the estimated elastic parameters decrease sequentially when using the traditional method, AG-based method, MSI method, and MMI method, respectively. This indicates a progressive improvement in the estimation accuracy of these four methods. By combining qualitative visualizations in the estimated 2D profiles and quantitative indicators of RMSE, the proposed MMI method demonstrates robust resistance to noise, further emphasizing its capability in addressing complex field cases with unknown challenges.

Fig. 10 shows the estimated V_P, V_S, and density obtained through different methods (the traditional method, the AG-based method and the MMI method) along with their true values at the 200th trace. From the 1D unlabeled trace illustration, it is evident that the MMI method achieves the best fit with the true values among the three. We omit the MSI results in this comparison, due to the minimal differences between the MSI and MMI results in the single-trace analysis.

Next, how the above-mentioned methods can gradually improve inversion accuracy will be discussed in detail. First, the multi-trace strategy provides more continuous results in the lateral direction. For a clear illustration, Fig. 11 compares the estimated V_P, V_S, and density obtained by the MSI and MMI methods when applied to noisy data with a SNR of 20 dB, and a local part from 0 to 150 ms is shown to highlight the details. The rectangles display the areas where the multi-trace method exhibits a notable improvement in horizontal continuity. The proposed MMI method emphasizes the significance of 2D convolution, as opposed to 1D convolution employed in the MSI method. The information from labeled traces can be inferred and incorporated into unlabeled traces through the 2D convolution operation, resulting in a closer connection between adjacent traces.

Second, the application of LPF to the PASD and EI is critical for improving lateral continuity and anti-noise capacity. Fig. 12 shows the influence of different low-pass frequencies on the final results obtained using the MSI method. Here, noisy data with a SNR of 5 dB is employed for testing. The results indicate that selecting a low-pass frequency within the range of 20–40 Hz achieves similar inversion results, confirming the robustness of the multi-attribute-assisted method with respect to filter selection. In this experiment, the optimal low-pass frequency is approximately set to 30 Hz. One principle for selecting the optimal low-pass frequency is to ensure it is higher than the dominant frequency of wavelet to avoid the loss of essential information.

Furthermore, the EI attribute and PASD provide more comprehensive information. To show the effectiveness of these two attributes, we conduct five comparative experiments as outlined in Table 2 for 5 dB noisy data. To avoid the impact caused by the network, all five comparative experiments uniformly adopt the single-trace network. The symbol LPF represents low pass filtering operation, INIT represents the initial models of V_P, V_S, and density,

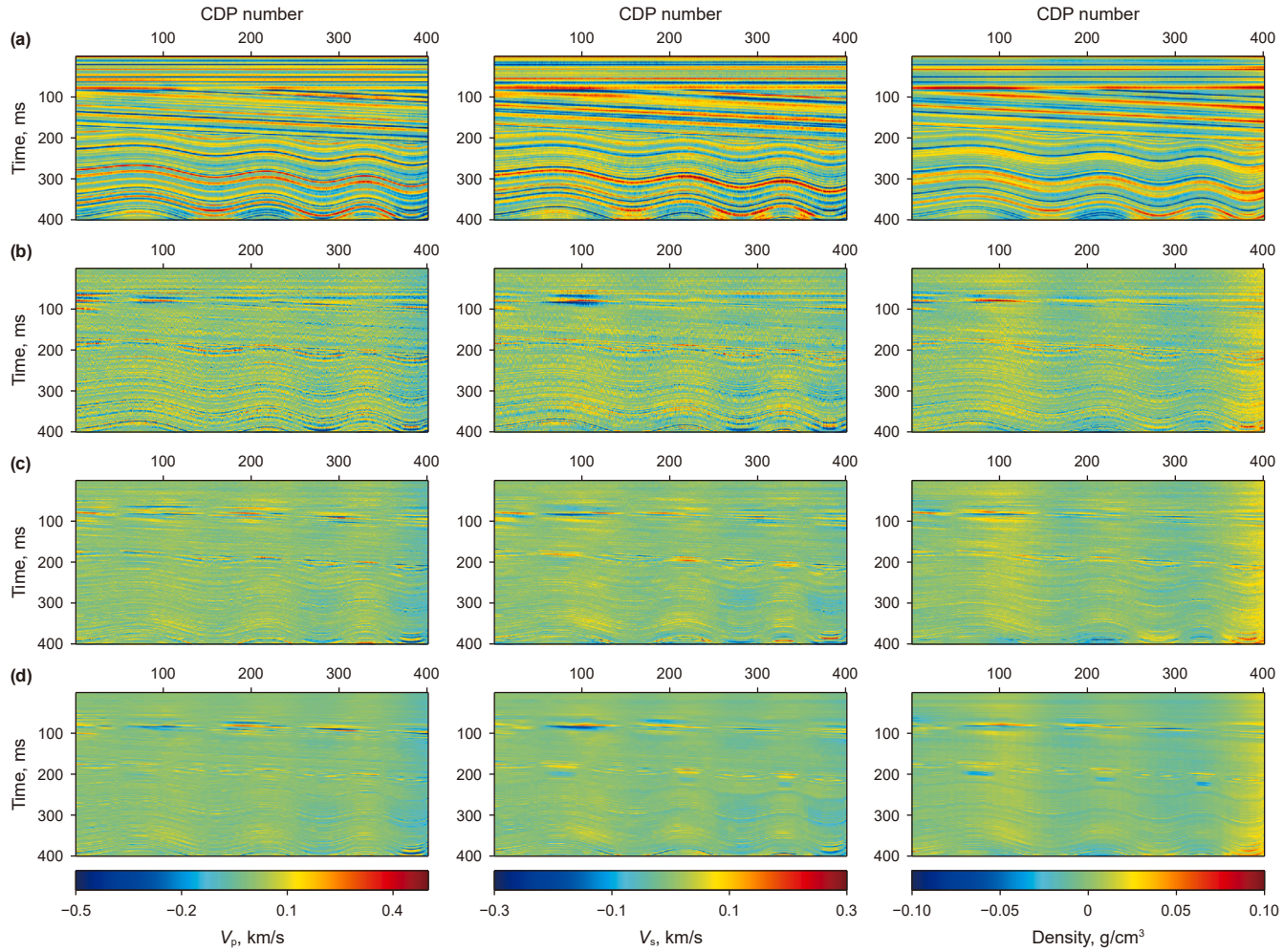


Fig. 9. The residuals between the estimated results in Fig. 8 and the true models. **(a)** to **(d)** represent the residuals using the traditional method, AG-based method, MSI method and MMI method, respectively. The first, second and third columns represent the residuals of V_P, V_S, and density, respectively.

Table 1
The RMSEs of V_P, V_S, and density using different methods. Bolds represent the best performing results in the comparisons. The same convention applies to all tables.

Method	P-wave	S-wave	Density
Traditional	0.1554	0.0958	0.0278
AG-based	0.0897	0.0490	0.0144
MSI	0.0585	0.0336	0.0112
MMI	**0.0475**	**0.0270**	**0.0090**

and + represents the concatenation operation along the channel dimension. Group 0 represents the MSI method with the designed multi-attributes as the input. In Group 1, the EI attribute in Group 0 is replaced with initial models of V_P, V_S, and density. The EI attribute and PASD are removed in Group 2 and Group 3, respectively. Lastly, the PAS is directly used as the input in Group 4.

The results reveal a decreasing trend for the estimation accuracy of V_P, V_S, and density from Group 0 to Group 4. Comparison between Group 0 and Group 1 demonstrates that the EI contains more information than the initial model. When compared with Group 0, Group 2 and Group 3 underscore the contributions of the filtered EI and PASD, respectively. It should be noted that the utilization of multiple attributes and the implementation of LPF are mutually dependent and play complementary roles. In the absence of the LPF operation, the presence of multiple attributes can lead to noise accumulation. In summary, the EI attribute offers more comprehensive information compared to the initial V_P, V_S, and density, and the PASD positively contributes to improving inversion accuracy. The utilization of multi-attributes has the potential to improve the model's generalization performance.

3.2. Field data example

3.2.1. Logging data example

In this part, we use pre-stack AG at 3 well logs (named W_1, W_2 and W_3) for testing. The quality of seismic data is high and the well-seismic calibration is reasonably accurate. The CDP numbers of W_1, W_2 and W_3 are 520, 91 and 382 in a 2D line. The central angles of PAS are 5°, 15° and 25°. Fig. 13 shows the PAS and its corresponding PASD and EI attributes of the 3 wells. To ensure robustness and mitigate selection bias, we conduct cross-validation experiments by using each well as a validation set while the others serve as training sets. Table 3 compares the Pearson correlation coefficient (PCC) and RMSE between the estimated V_P, V_S, density and the labels using the traditional method, PAS-based method and MSI method. Fig. 14 shows the obtained V_P, V_S, and density when W_3 serves as the validation set using the above three methods.

From Table 3 and Fig. 14, we can see that the MSI method outperforms the PAS-based method and traditional method.

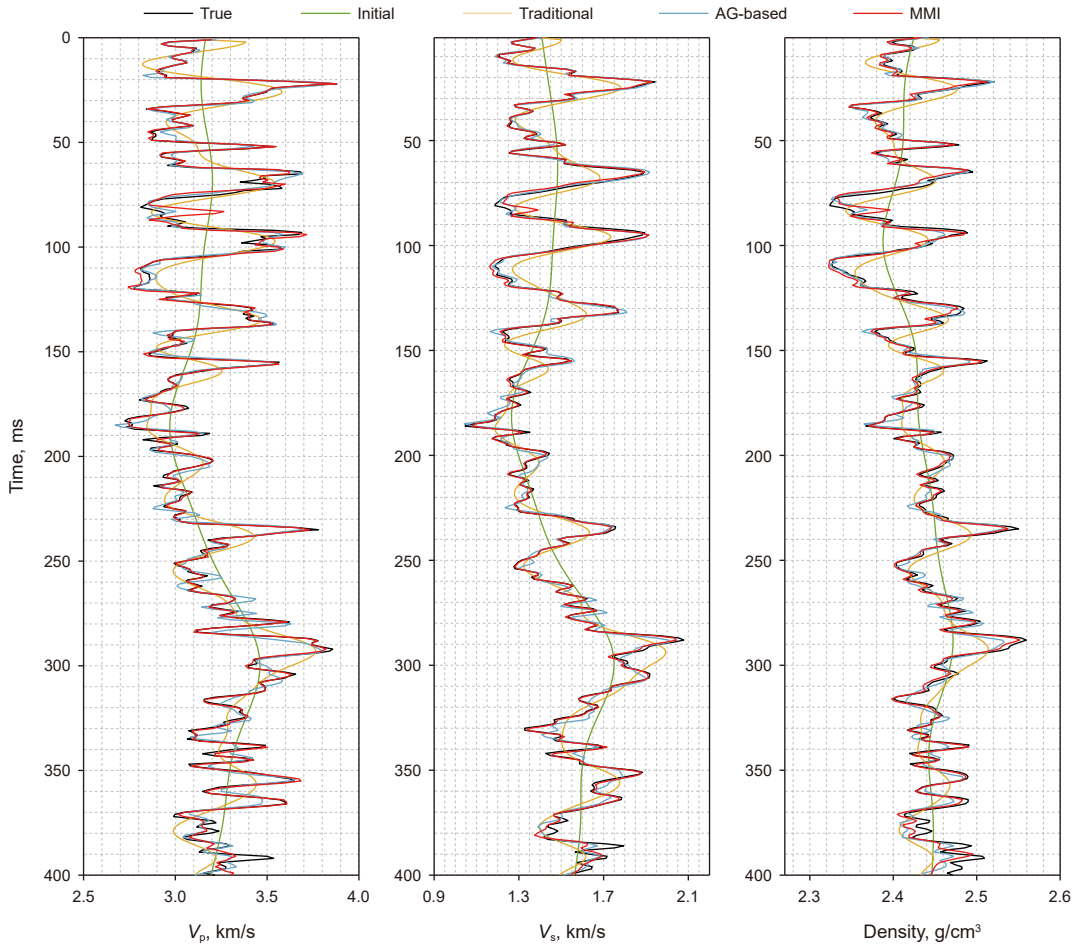


Fig. 10. Comparisons between inversion results for V_P, V_S, and density obtained through different methods and their true values at the 200th trace.

Because of high quality of the data, the traditional method obtains satisfactory estimated results, but the accuracy is still open to improvement, especially for density. The high precision inversion of density using the traditional method is challenging because V_P has a noticeable impact on density inversion. The estimated density for W_3 using PAS-based method is slightly improved compared with the traditional method, but it is contrary to the logging data somewhere (such as near 1700 ms in Fig. 14). The estimated results using the MSI method tend to get more accurate values, which demonstrates multi-attributes can enhance the model's generalization performance.

3.2.2. *Two-dimensional field data example*

We apply the proposed method to 2D field data in a China basin. The time range is 3600–4400 ms with a time interval of 2 ms, and the range of CDPs is from 200 to 1400. In each CDP, there are 20 angles (from 2° to 40° with an interval of 2°). We select 3 groups of angles: 2–14°, 16–28° and 30–40° to calculate PAS, shown in Fig. 15(a). The central angles of the three PAS are 8°, 22° and 35°, respectively. From the 2D field data, three horizons are picked: H_{60}, H_{70}, and H_{72}. There are two well-logs located at CDP = 459 and 1260, named A and B, respectively. The initial models of V_P, V_S, and density are constructed via horizon-guided well-log interpolation, as shown in Fig. 15(b). We only show the profile between H_{60} and H_{70} for clear display.

We choose well A to build the training set, and well B for blind testing. It is worth noting that the high-frequency compositions within well-logging data may deeply affect the inversion results, and it is smoothed prior to training.

Fig. 16(a)–(d) show the estimated V_P, V_S, and density using the traditional method, AG-based method, MSI method, and MMI method. The features of the obtained results are similar to those of the synthetic example, i.e., DL methods exhibit higher resolution compared to traditional method. Noticeable spot-like results are obtained using AG-based method, which is probably due to the impact of noise in the prestack AG. Compared with AG-based method, MSI method improves the horizontal continuity, describes thin layers more clearly, and suppresses spot-like noise meanwhile. The result of MMI is similar to that of MSI in terms of profile, but the MMI method provides enhanced details in localized areas. For example, details inside the elliptical region appear more distinct, and thin layers indicated by arrows are more focused.

Fig. 17 presents the estimated V_P, V_S, and density of the test blind well. To clearly depict the trends, low-pass filtering (0-0-60–80 Hz) is applied to the true well values. Compared to the traditional method, the MMI method shows closer proximity to the true values at the peak, such as the area around 4300 ms.

Table 4 quantitatively characterizes the inversion results of the blind well by calculating the RMSE and PCC between the inversion results and the true well-logging data. It can be observed that the MMI method consistently outperforms the MSI method, AG-based method, and traditional method in terms of accuracy, which demonstrates the effectiveness of the proposed method.

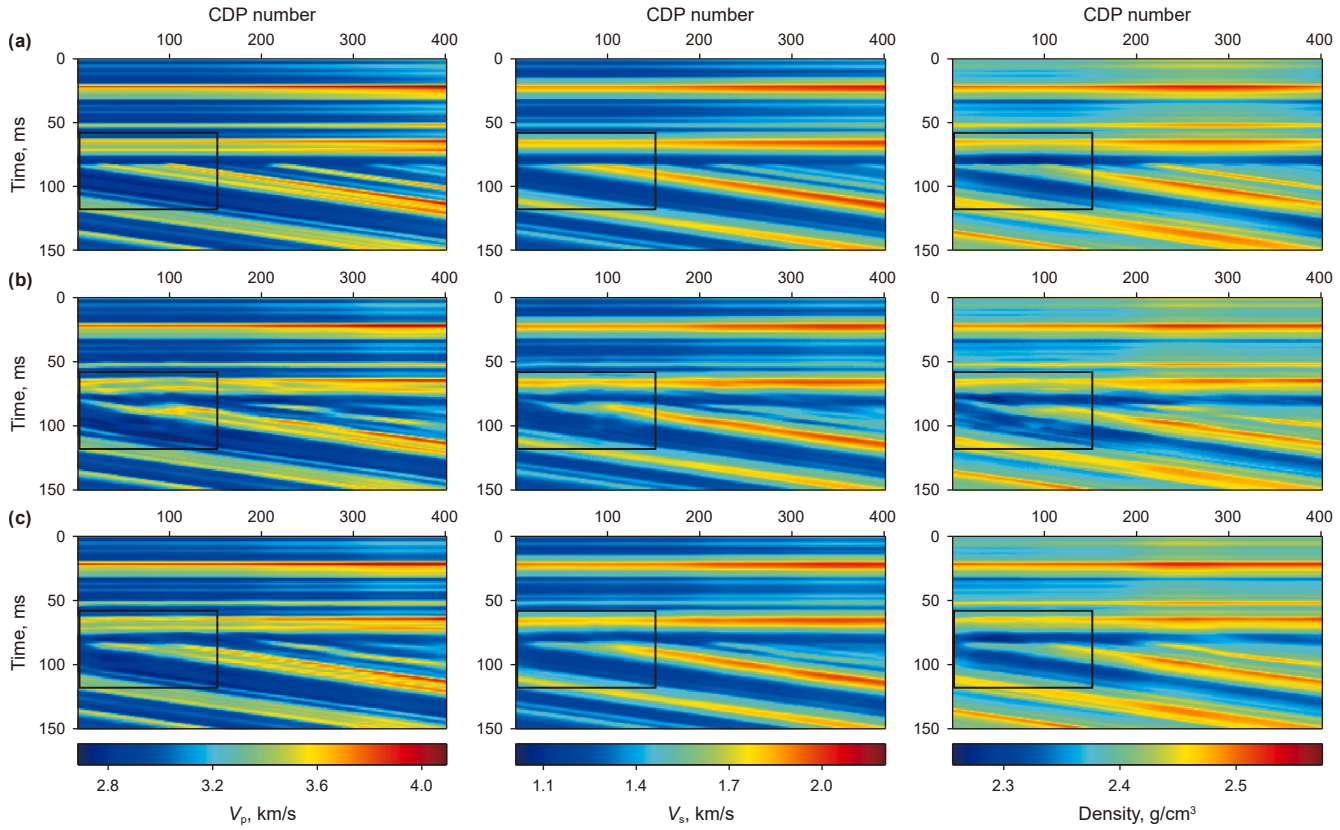


Fig. 11. Comparisons of the estimated results in the 0-150 ms part when applied to noisy seismic data with SNRs of 20 dB. **(a)** to **(c)** represent the true models and the estimated results using the MSI and MMI method. The first, second and third columns represent V_P, V_S, and density, respectively.

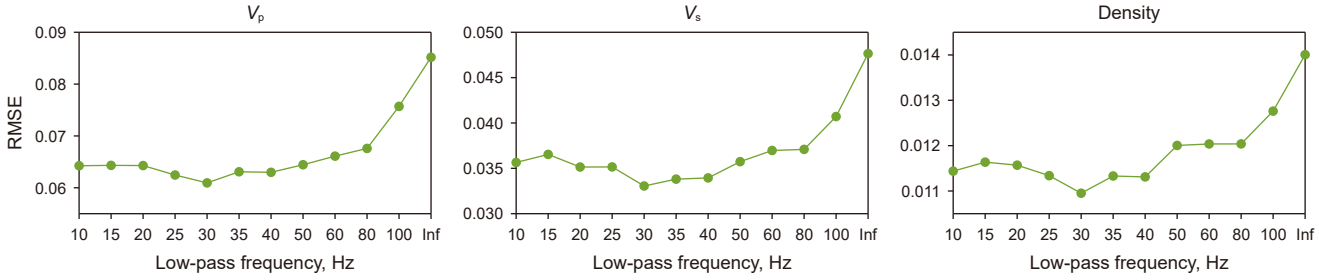


Fig. 12. The influence of low-pass frequency on the estimated results in terms of RMSE. The results are obtained using the MSI method.

Table 2
Components of input data and RMSEs between estimated results and true model. All groups uniformly adopt the single-trace network.

Group	The components of input data	P-wave	S-wave	Density
0	PAS + LPF(PASD) + LPF(EI)	**0.0585**	**0.0336**	**0.0112**
1	PAS + LPF(PASD) + INIT	0.0610	0.0349	0.0116
2	PAS + LPF(PASD)	0.0639	0.0360	0.0116
3	PAS + LPF(EI)	0.0690	0.0376	0.0124
4	PAS	0.0791	0.0442	0.0132

4. Discussion

4.1. Sensitivity analysis

To analyze the robustness and applicability of the proposed method, we first compare the inversion results of noisy AG with different SNRs. Specifically, four different methods are compared: MMI, MSI, PAS-based and AG-based algorithms. The parameter of a in the modified tanh activation function is fixed to 0.2. Fig. 18 shows the inversion results in noisy cases with different SNRs via different methods. We average the results of five independent runs to ensure reliability of the inversion results.

It can be observed from Fig. 18 that the RMSE exhibits an increasing trend as the SNR decreases. The RMSEs of the AG-based, PAS-based, MSI, and MMI methods decrease successively with a specific SNR, indicating an upward tendency in the parameter estimation accuracy. The RMSE difference between SNR = 2 and SNR = 20 can be used to measure the anti-noise performance of different methods. The anti-noise performance of the AG-based, PAS-based, and multi-attribute methods increases successively. It suggests that the partial stack strategy can effectively mitigate the influence of noise, and the proposed MMI method exhibits the best anti-noise performance.

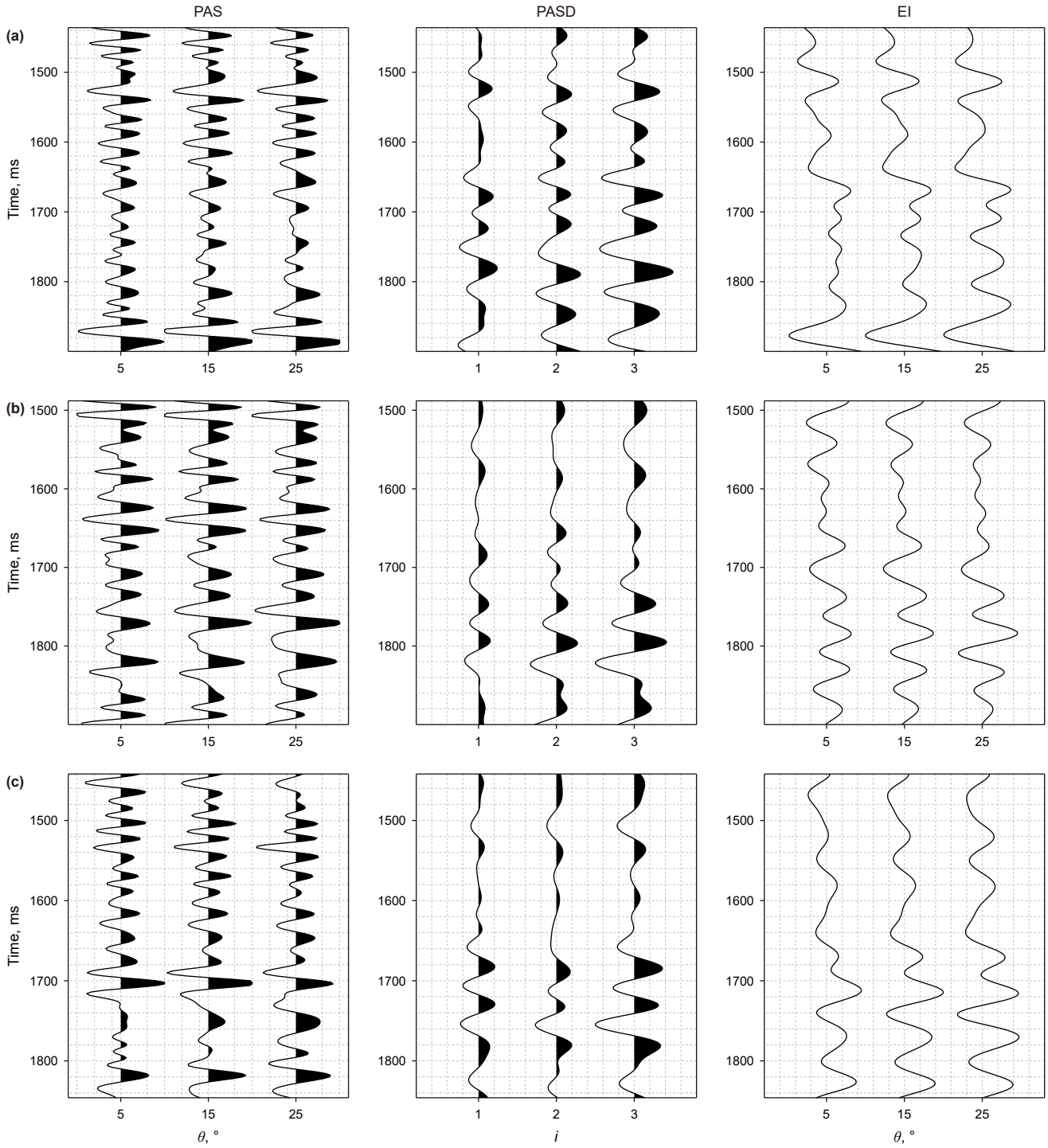


Fig. 13. The network input. **(a)**, **(b)** and **(c)** are of W_1, W_2 and W_3. The first to third columns represent the PAS, PASD and EI attributes. The PASD and EI are the results after low pass filtering.

Next, the role of the modified tanh activation function is discussed. Fig. 19 quantifies the inversion results of noisy AG with different SNRs using the MSI method when the parameter a in the modified tanh activation function equals 0.1, 0.2, 0.4, 0.6, 0.8 and 1. The fixed hyperparameters of the network are utilized for a fair comparative analysis. To assess the reliability of the estimated results, the average of the results from five independent experiments is taken as the final result. As mentioned above, the value of $a_{\min}$ equals 0.1722 and it is a reasonable constraint when a is larger than $a_{\min}$. Fig. 19 indicates that when $a = 0.1$, the obtained RMSE increases significantly due to an improper constraint. When $a \geq 0.2$, as the value of a increases, the RMSE also increases. This experiment emphasizes the importance of selecting an appropriate range for the parameter a. In general, the modified activation function with an appropriate parameter serves as an effective constraint in noisy cases.

Table 3
Comparisons of PCC and RMSE between the estimated V_P, V_S, and density and the true values. The results of PAS-based method and MSI method are of validation wells in cross-validation experiments.

Parameters		Traditional		PAS-based		MSI	
		PCC	RMSE	PCC	RMSE	PCC	RMSE
W_1	V_P	0.7584	0.2327	0.7788	0.2062	**0.8025**	**0.1945**
	V_S	0.8040	0.1783	0.8105	0.1720	**0.8312**	**0.1618**
	Density	0.7553	0.0512	0.8059	0.0378	**0.8264**	**0.0376**
W_2	V_P	0.7924	0.2936	0.7827	0.2297	**0.8504**	**0.1879**
	V_S	0.7976	0.2014	0.8005	0.1747	**0.8343**	**0.1481**
	Density	0.7543	0.0573	0.7514	0.0562	**0.7732**	**0.0542**
W_3	V_P	0.7975	0.3375	0.8061	0.2318	**0.8848**	**0.1803**
	V_S	0.7765	0.2278	0.7741	0.1906	**0.8268**	**0.1692**
	Density	0.6727	0.0789	0.7201	0.0564	**0.8036**	**0.0459**

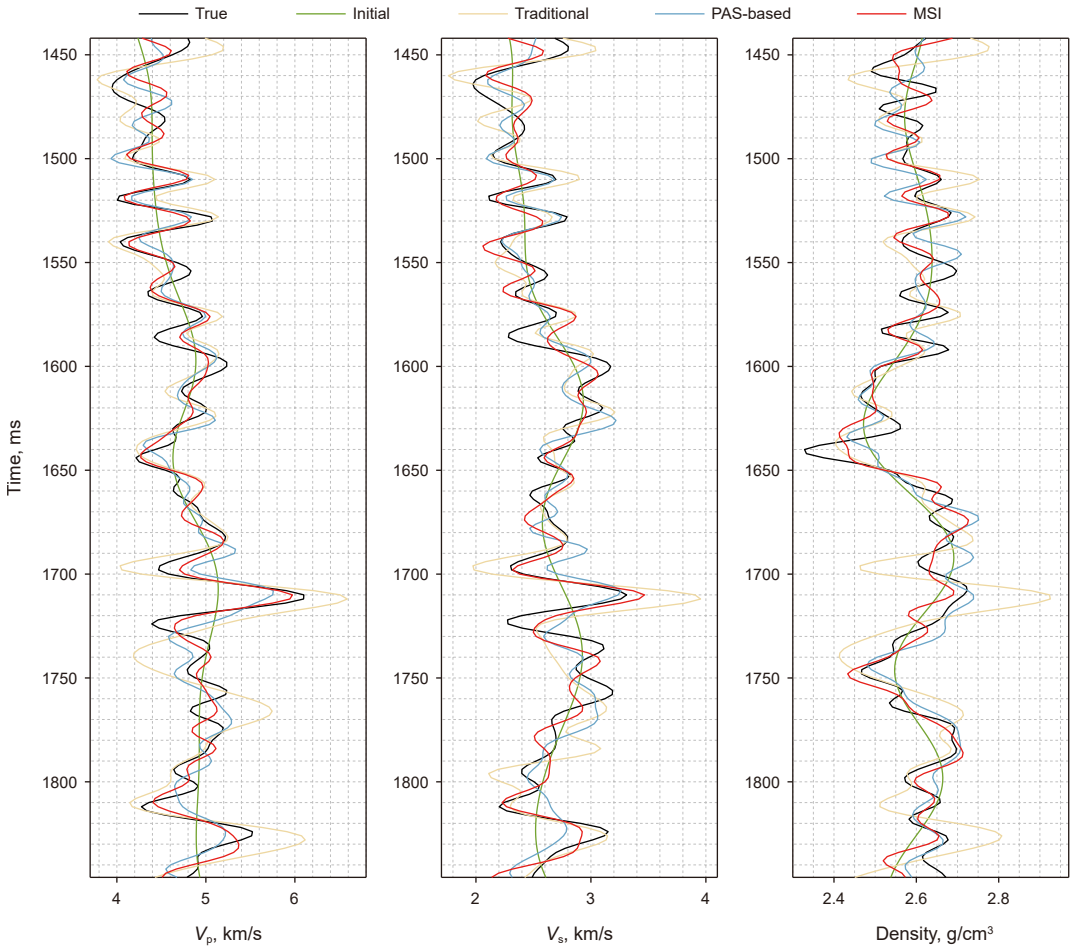


Fig. 14. The true and estimated V_P, V_S, and density of W_3.

4.2. Cross-validation for hyperparameter optimization

Neural networks have a large number of hyperparameters, which significantly impact the final performance. Cross-validation is a common tool to optimize parameter combinations, during which one well is selected for validation while the remaining wells are used for training. For V_P, we calculate the RMSE between the estimated and true values of the selected validation well. After changing validation well in turn, the average RMSE across all validation wells is used to assess the selected parameters. For different parameter combinations, a series of average RMSEs $\mathbf{R} = \{R_1, R_2, \ldots, R_P\}$ are obtained, where P represents the number of tested parameter combinations. We define the score $\mathbf{S}$ for each parameter combination as:

$$\mathbf{S} = \sigma(-Z_{\text{score}}(\mathbf{R})), \tag{20}$$

where $Z_{\text{score}}(\cdot)$ denotes Z-score normalization, and $\sigma(\cdot)$ denotes the Softmax function, which is expressed as:

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}. \tag{21}$$

Here, Softmax operation amplifies differences between parameter combinations. We further normalize the scores using Eq. (22):

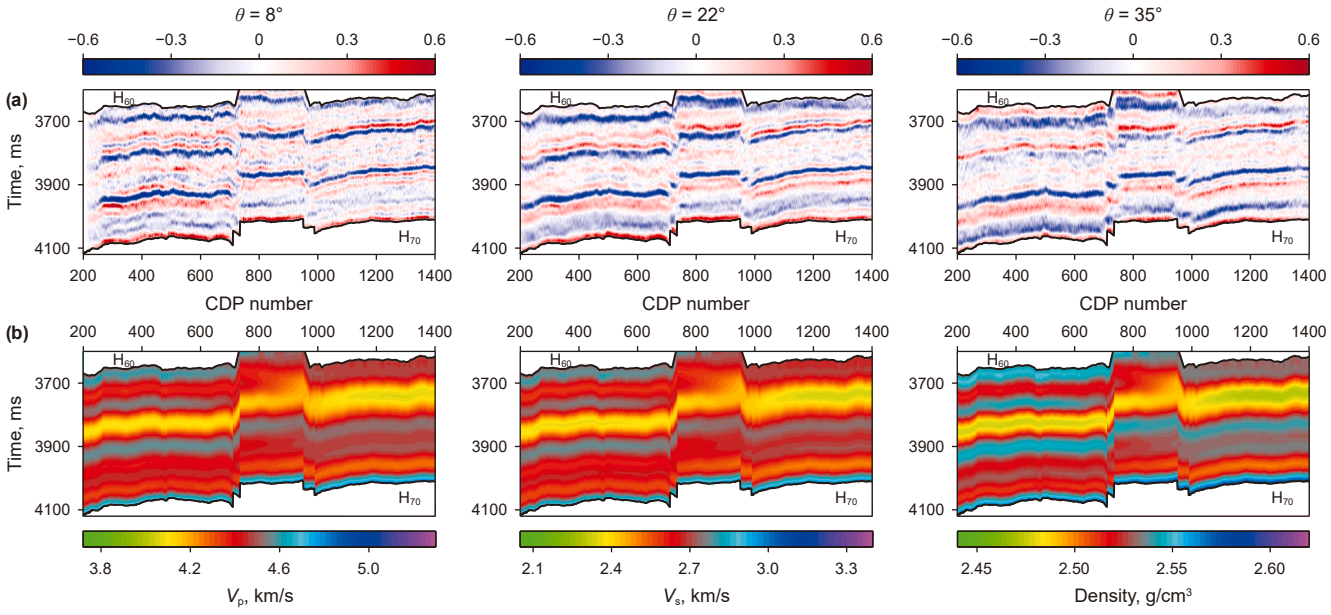


Fig. 15. The **(a)** PAS and **(b)** initial models of V_P, V_S, and density.

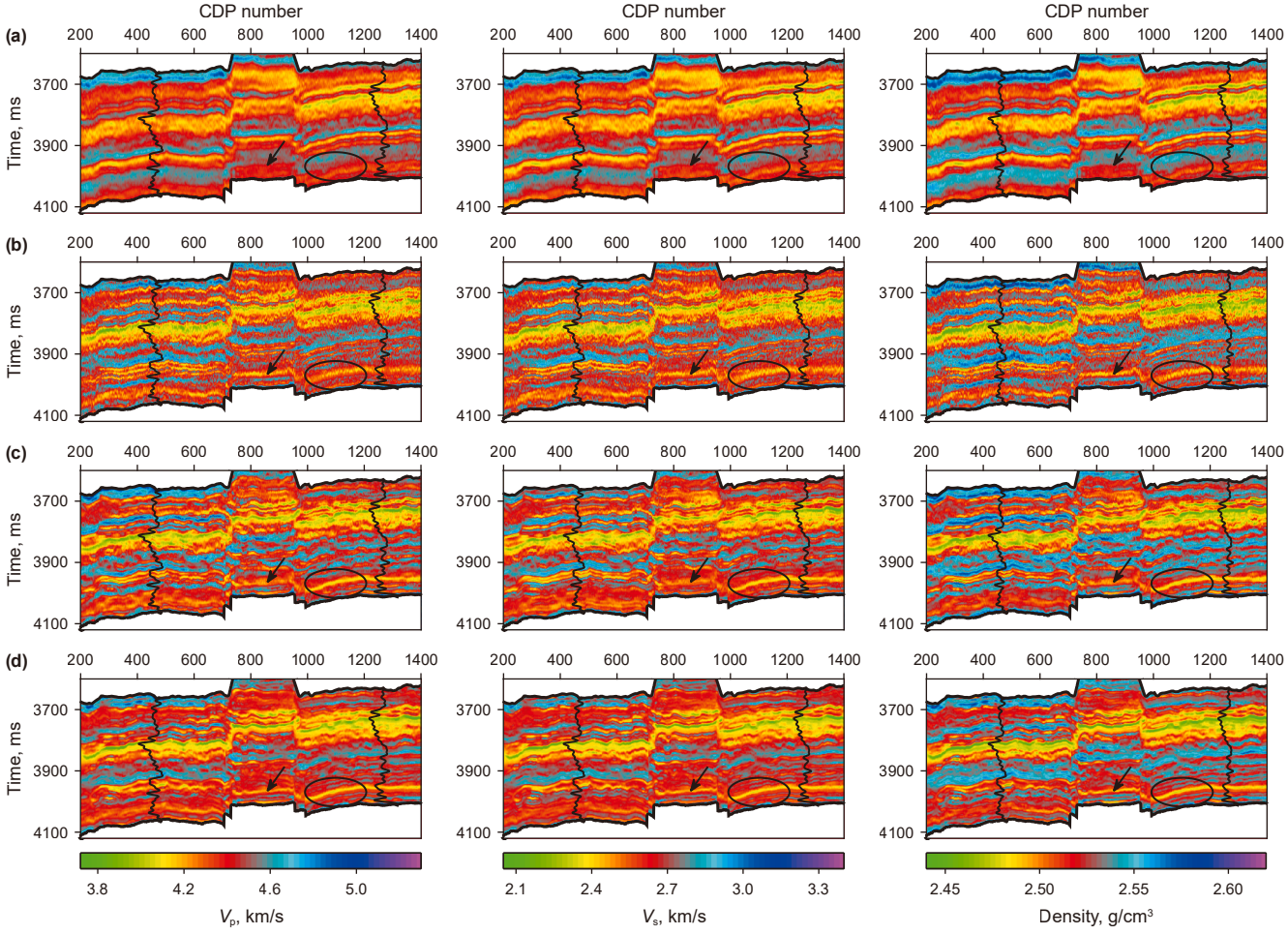


Fig. 16. Comparisons of estimated elastic parameters using different methods. **(a)** to **(d)** are the estimated results using the traditional method, AG-based method, MSI method and MMI method, respectively. The first to third columns represent V_P, V_S, and density, respectively. The black curves denote two wells.

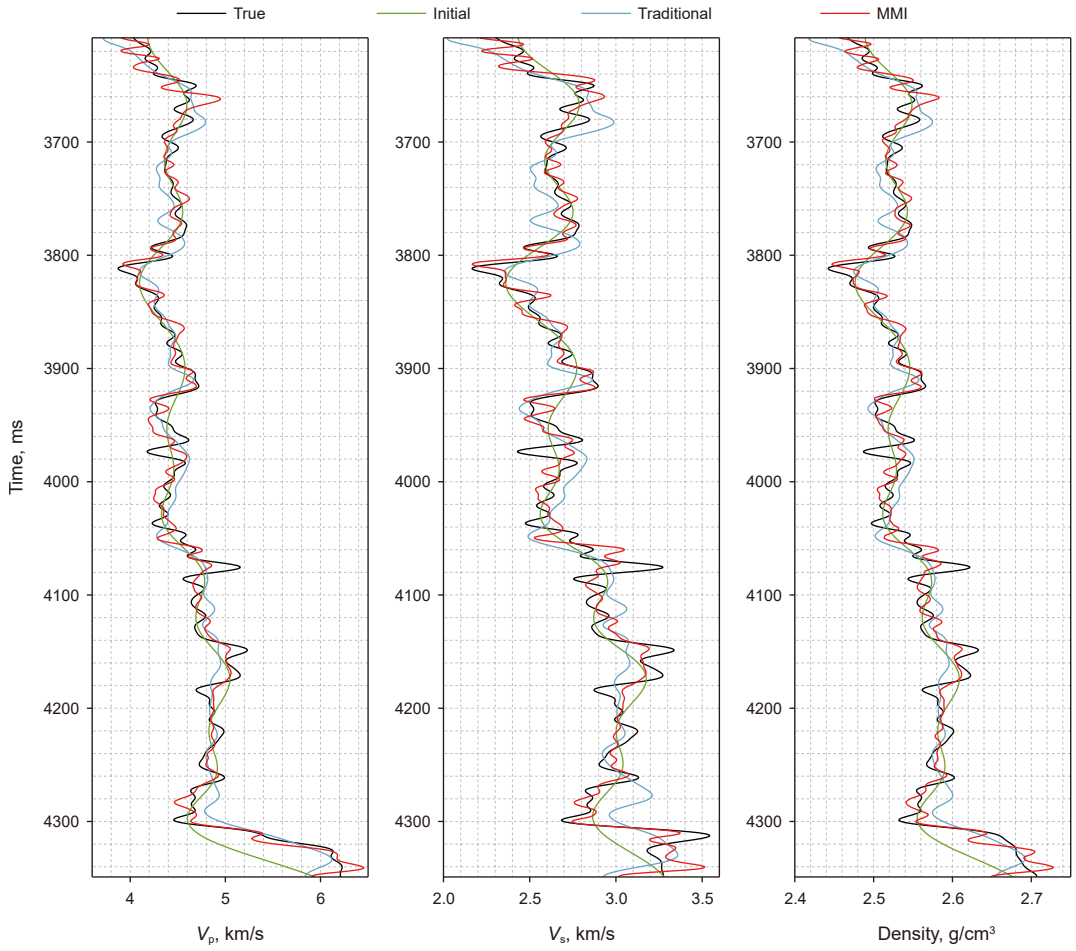


Fig. 17. The true and estimated V_P, V_S, and density of the blind well.

Table 4
The calculated PCC and RMSE of V_P, V_S, and density using different methods at the blind well location of 2D field data.

Method	Index	P-wave	S-wave	Density
Traditional method	PCC	0.9329	0.8396	0.9037
	RMSE	0.1504	0.1430	0.0207
AG-based method	PCC	0.9334	0.8744	0.9052
	RMSE	0.1576	0.1319	0.0216
MSI method	PCC	0.9443	0.9063	0.9320
	RMSE	0.1452	0.1133	0.0180
MMI method	PCC	**0.9589**	**0.9242**	**0.9451**
	RMSE	**0.1201**	**0.0981**	**0.0157**

$$\mathbf{S}_{\text{norm}} = 100\frac{\mathbf{S} - \min(\mathbf{S})}{\max(\mathbf{S}) - \min(\mathbf{S})}. \tag{22}$$

According to Eqs. (20) and (22), the scores are converted to a range between 0 and 100. The higher scores, the better parameter combinations. Similarly, scores for V_S and density can be obtained, and these three scores can be averaged to determine the overall scores for the parameter combinations.

Fig. 20 shows the cross-validation scores of two key hyperparameters (K and C) in the MMI network. Fig. 20(a)–(c) indicates the scores of V_P, V_S, and density, respectively, and the overall average score is shown in Fig. 20(d). The final scores show that the optimal parameter combination is $K = 21$ and $C = 24$, which is employed in the synthetic data example. It is reasonable to choose a convolution kernel size of 21×3 because we use long strips of patches. Further increasing K may result in higher computational requirements and the risk of getting trapped in local minima. To improve network efficiency, dilated convolution may be considered in the further work. The value of C controls network complexity, and a value that is too large may result in an excessive number of network parameters and overfitting issues during training.

4.3. The advantages and limitations of elastic parameter inversion using AVO and DL

The AVO inversion based on the Zoeppritz equation and its approximations is a traditional and classical theory. The Zoeppritz equation simplifies the complex relationship between seismic amplitude and subsurface properties, making elastic parameter inversion theoretically feasible. Inversion based on AVO theory offers advantages in interpretability and stability but faces challenges in terms of accuracy. The Zoeppritz equation is based on the plane-wave assumption, and its practical applicability is limited because field data often deviate from this

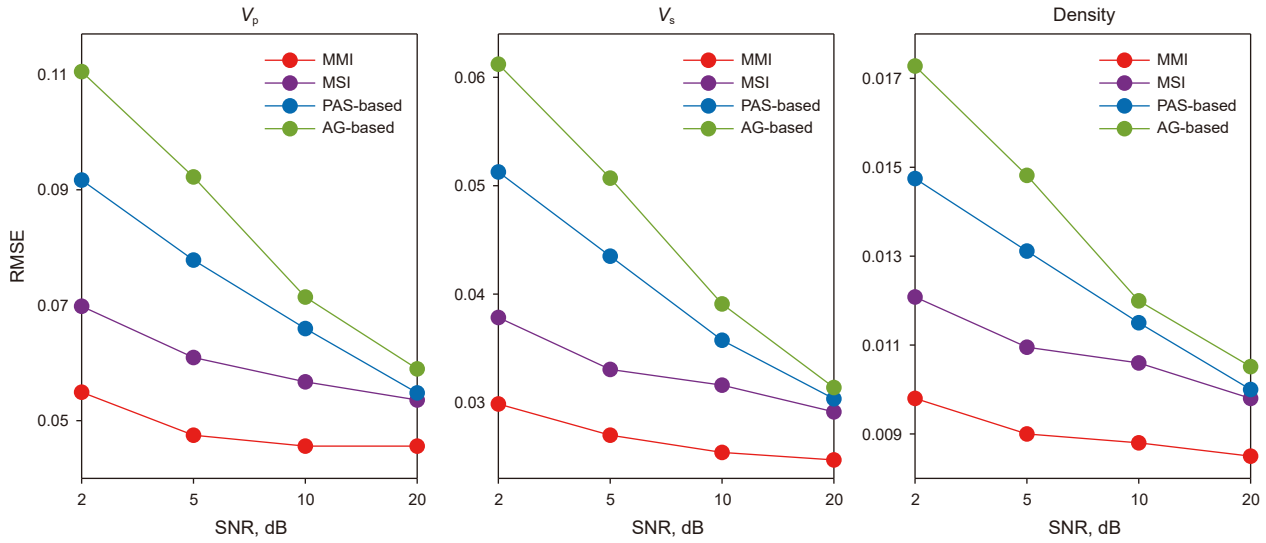


Fig. 18. The RMSEs of the estimated results in noisy cases with different SNRs via different methods ($a = 0.2$ in modified-tanh).

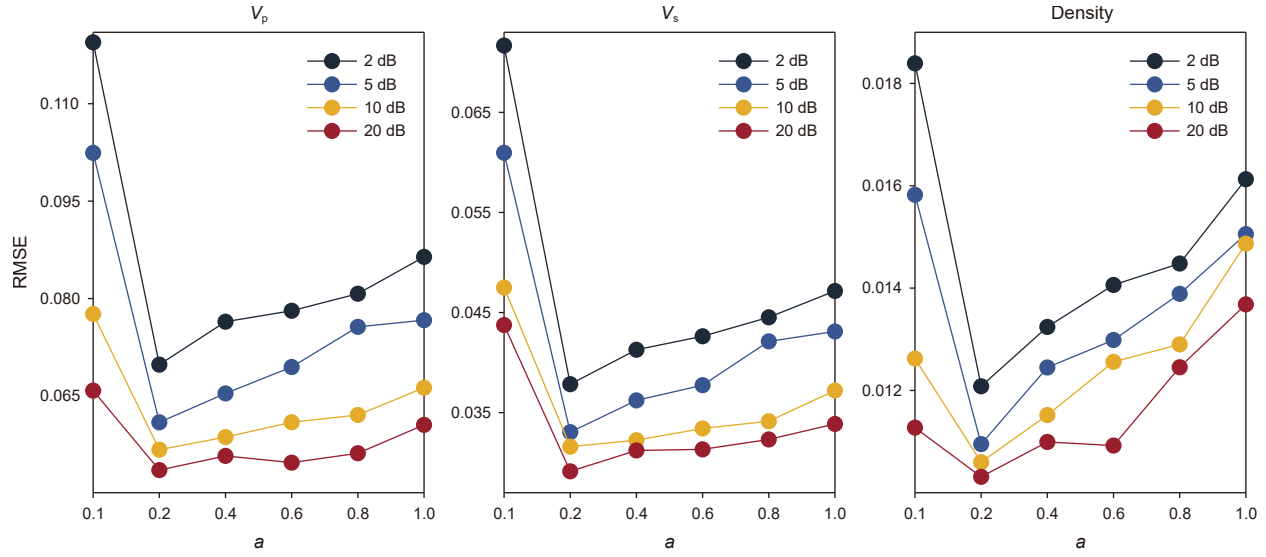


Fig. 19. The RMSEs of the estimated results for noisy AG with different SNRs using MSI method when the parameter a takes the values 0.1, 0.2, 0.4, 0.6, 0.8 and 1.

assumption, especially in cases involving anisotropy, inhomogeneity, and inelasticity. These deviations can lead to significant mismatches (particularly for large-angle seismic data) during well-seismic calibration when forward modeling relies on the plane-wave assumption.

Deep learning-based elastic parameter inversion is a nonlinear mapping method that effectively learns the relationship between seismic data and elastic parameter labels. The representativeness of these labels significantly impacts inversion accuracy. When dealing with relatively simple geological structures, such as primarily horizontally layered media, a small number of logging labels can suffice to build a representative training set, explaining why the proposed method achieves high precision in elastic parameter inversion. However, for complex cases, the proposed method may overfit and fail to accurately capture abrupt geobodies, particularly when low-frequency models contain significant errors or structural biases. Additional logging labels and rational physical constraints are needed for assistance.

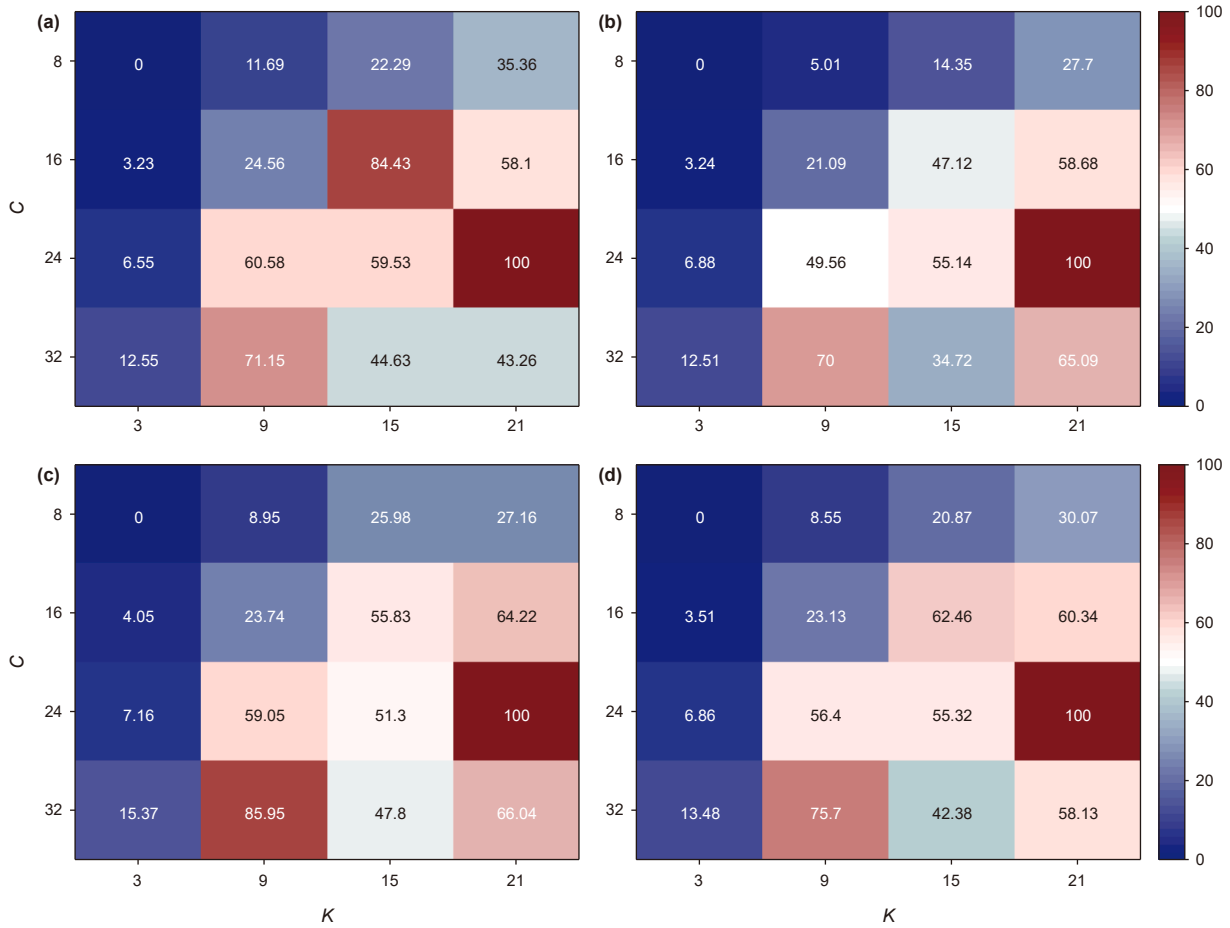


Fig. 20. The scores of hyper-parameters (K and C) using cross validation. **(a)**, **(b)** and **(c)** are the scores of V_P, V_S, and density; **(d)** The average of **(a)**, **(b)** and **(c)** as overall scores.

5. Conclusion

We propose a multi-attribute-assisted supervised deep learning algorithm for elastic parameter inversion. The multiple attributes include partial angle stacks (PAS), partial angle stack differences (PASD) and elastic impedance (EI). To address the limitations of single-trace inversion algorithms, a multi-trace inversion method is designed using a 2D convolution operator during which patching and masking strategies are used to augment single-trace labels. Comparative experiments demonstrate that the EI attribute offers richer information compared to the initial model. Additionally, the PASD attribute positively enhances the inversion accuracy. The use of a low-pass filter operation for the PASD and EI attributes and a modified tanh activation function for the network also help improve the estimation accuracy of elastic parameters. Synthetic data examples and field data applications confirm that the multi-attribute-assisted multi-trace inversion method achieves high accuracy, robust noise resistance and satisfactory lateral continuity.

CRediT authorship contribution statement

Ren Luo: Writing – original draft, Software, Methodology, Formal analysis. **Huai-Zhen Chen:** Writing – review & editing, Validation, Investigation, Data curation. **Jian-Hua Geng:** Writing – review & editing, Funding acquisition, Data curation. **Ben-Feng Wang:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (42330805, 42374134), the China National Petroleum Corporation (CNPC) Innovation Found (2022DQ02-0612), the Science and Technology Commission of Shanghai Municipality (23JC1400502) and the Fundamental Research Funds for the Central Universities.

References

Aki, K., Richards, P.G., 1980. Quantitative Seismology: Theory and Methods. W. H. Freeman.

Barens, L., Hansen, H.J., 2006. Pre-stack seismic inversion of offset stacked data: diagnostics and remedies for overestimated angles of incidence, a case study. In: SEG Technical Program Expanded Abstracts, pp. 1650–1654, 2006.

Biswas, R., Sen, M.K., Das, V., et al., 2019. Prestack and poststack inversion using a physics-guided convolutional neural network. Interpretation 7 (3), SE161–SE174. https://doi.org/10.1190/INT-2018-0236.1.

Bosch, M., Mukerji, T., Gonzalez, E.F., 2010. Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: a review. Geophysics 75 (5), 75A165–175A176. https://doi.org/10.1190/1.3478209.

Cao, D., An, P., Liu, S., 2018. Elastic-parameters inversion from EI based on the deep-learning method. In: SEG Technical Program Expanded Abstracts, pp. 640–644, 2018.

Chen, F., Zong, Z., Yang, Y., et al., 2022. Amplitude-variation-with-offset inversion using P- to S-wave velocity ratio and P-wave velocity. Geophysics 87 (4), N63–N74. https://doi.org/10.1190/geo2021-0623.1.
Connolly, P., 1999. Elastic impedance. Lead. Edge 18 (4), 438.
Dai, R., Yin, C., Peng, D., 2022. AVA inversion of pre-stack seismic data with cross-gradient constraints. J. Appl. Geophys. 199, 104594. https://doi.org/10.1016/j.jappgeo.2022.104594.
Di, H., Li, Z., Abubakar, A., 2022. Using relative geologic time to constrain convolutional neural network-based seismic interpretation and property estimation. Geophysics 87 (2), IM25–IM35. https://doi.org/10.1190/geo2021-0257.1.
Dunham, M.W., 2021. Are seismic attributes still helpful for deep learning?. In: First International Meeting for Applied Geoscience & Energy Expanded Abstracts, pp. 1561–1565.
Fatti, J.L., Smith, G.C., Vail, P.J., et al., 1994. Detection of gas in sandstone reservoirs using AVO analysis: a 3-D seismic case history using the geostack technique. Geophysics 59 (9), 1362–1376. https://doi.org/10.1190/1.1443695.
Goodway, B., Chen, T., Downton, J., 1997. Improved AVO fluid detection and lithology discrimination using Lamé petrophysical parameters; "λρ,""μρ,"&"λ/μ fluid stack," from P and S inversions. In: SEG Technical Program Expanded Abstracts, pp. 183–186, 1997.
Goyal, P., Dollár, P., Girshick, R., et al., 2018. Accurate, Large Minibatch SGD: Training Imagenet in 1 Hour. https://doi.org/10.48550/arXiv.1706.02677.
Hampson, D.P., Russell, B.H., Bankhead, B., 2005. SImultaneous inversion of pre-stack seismic data. In: SEG Technical Program Expanded Abstracts, pp. 1633–1637, 2005.
Han, D.h., Nur, A., Morgan, D., 1986. Effects of porosity and clay content on wave velocities in sandstones. Geophysics 51 (11), 2093–2107. https://doi.org/10.1190/1.1442062.
Li, X., Pan, X., Liu, H., et al., 2025. Seismic inversion for fracture properties using model-data dual-driven network. IEEE Trans. Geosci. Rem. Sens. 63, 1–9. https://doi.org/10.1109/tgrs.2025.3591562.
Luo, R., Wang, B., Chen, H., et al., 2023. Deep learning prestack seismic inversion based on multi-seismic features. In: 84th EAGE Annual Conference & Exhibition, pp. 1–5.
Ma, J., Geng, J., Guo, T., 2011. Using linearized Bayesian method to extract elastic parameters from elastic impedance. In: SEG Technical Program Expanded Abstracts, pp. 2762–2766, 2011.
Mallick, S., 2001. AVO and elastic impedance. Lead. Edge 20 (10), 1094–1104. https://doi.org/10.1190/1.1487239.
Ostrander, W.J., 1984. Plane-wave reflection coefficients for gas sands at non-normal angles of incidence. Geophysics 49 (10), 1637. https://doi.org/10.1190/1.1441571.
Russell, B.H., 1988. 8. Part 8 - model-based Inversion. In: Introduction to Seismic Inversion Methods, pp. 8-1-8–17.
Shalabi, L.A., Shaaban, Z., Kasasbeh, B., 2006. Data mining: a preprocessing engine. Applied Science University 2 (9), 735–739. https://doi.org/10.3844/jcssp.2006.735.739.
Shuey, R.T., 1985. A simplification of the Zoeppritz equations. Geophysics 50 (4), 609–614. https://doi.org/10.1190/1.1441936.
Smith, G.C., Gidlow, P.M., 1987. Weighted stacking for rock property estimation and detection of gas. Geophys. Prospect. 35 (9), 993–1014. https://doi.org/10.1111/j.1365-2478.1987.tb00856.x.
Song, L., Yin, X., Zong, Z., et al., 2022. Cooperative multinetworks semi-supervised pre-stack seismic inversion. Geophys. J. Int. 230 (3), 1878–1894.
Sun, S., Zhao, L., Chen, H., et al., 2023. Prestack seismic inversion for elastic parameters using model-data-driven generative adversarial networks. Geophysics 88 (2), M87–M103. https://doi.org/10.1190/geo2022-0314.1.
Wang, B., 2005. Elastic impedance inversion and its application. Prog. Geophys. 20 (1), 89–92.
Wang, L., Meng, D., Wu, B., 2021a. Seismic inversion via closed-loop fully convolutional residual network and transfer learning. Geophysics 86 (5), R671–R683. https://doi.org/10.1190/geo2020-0297.1.
Wang, P., Chen, X., Li, J., et al., 2021b. Lateral constrained prestack seismic inversion based on difference angle gathers. IEEE Geosci. Remote Sens. Lett. 18 (12), 2177–2181.
Wang, Y., Ge, Q., Lu, W., et al., 2020. Well-Logging constrained seismic inversion based on closed-loop convolutional neural network. IEEE Geosci. Remote Sens. Lett. 58 (8), 5564–5574.
Wang, Y., Wang, Q., Lu, W., et al., 2022. Physics-Constrained seismic impedance inversion based on deep learning. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/lgrs.2021.3072132.
Wang, Z., Gao, J., Zhao, H., et al., 2023. Three-parameter prestack nonlinear inversion constrained by gradient structure similarity. Geophysics 88 (1), N1–N19. https://doi.org/10.1190/geo2021-0311.1.
Whitcombe, D.N., 2002. Elastic impedance normalization. Geophysics 67 (1), 60–62. https://doi.org/10.1190/1.1451331.
Wu, X., Ma, J., Si, X., et al., 2023. Sensing prior constraints in deep neural networks for solving exploration geophysical problems. Proc. Natl. Acad. Sci. 120 (23), e2219573120. https://doi.org/10.1073/pnas.2219573120.
Wu, X., Yan, S., Bi, Z., et al., 2021. Deep learning for multidimensional seismic impedance inversion. Geophysics 86 (5), R735–R745. https://doi.org/10.1190/geo2020-0564.1.
Xu, M., Zhao, L., Gao, S., et al., 2022. Joint use of multiseismic information for lithofacies prediction via supervised convolutional neural networks. Geophysics 87 (5), M151–M162. https://doi.org/10.1190/geo2021-0554.1.
Zhang, J., Sun, H., Huang, X.-G., et al., 2026. Joint physics and model-guided prestack seismic inversion using double dual neural network. Pet. Sci. 23 (1), 157–177. https://doi.org/10.1016/j.petsci.2025.09.008.
Zhang, J., Zhao, X., Chen, Y., et al., 2023. Domain knowledge-guided data-driven prestack seismic inversion using deep learning. Geophysics 88 (2), M31–M47. https://doi.org/10.1190/geo2021-0560.1.
Zhao, L., Meng, J., Xu, M., et al., 2026. Seismic lithofacies prediction via deep learning with prior geological constraints. Geophysics 91 (1), MR33–MR50. https://doi.org/10.1190/geo2024-0800.
Zhong, Z., Carr, T.R., Wu, X., et al., 2019. Application of a convolutional neural network in permeability prediction: a case study in the Jacksonburg-Stringtown oil field, West Virginia, USA. Geophysics 84 (6), B363–B373. https://doi.org/10.1190/geo2018-0588.1.
Zong, Z., Yin, X., Wu, G., 2012. Elastic impedance variation with angle inversion for elastic parameters. J. Geophys. Eng. 9 (3), 247–260. https://doi.org/10.1088/1742-2132/9/3/247.
Zou, C., Zhao, L., Hong, F., et al., 2023. A comparison of machine-learning methods to predict porosity in carbonate reservoirs from seismic-derived elastic properties. Geophysics 88 (2), B101–B120. https://doi.org/10.1190/geo2021-0342.1.