Petroleum Science 23 (2026) 3972–3987

Contents lists available at ScienceDirect

Petroleum Science

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

Original Paper

Full waveform inversion via BEEMD-based gradient decomposition

Yi-Lin Lu [a], Jian-Ping Huang [a,*], Liang Chen [b], Guo-Long Li [a], Qiu-Yang Wang [a]

[a] *Geosciences Department, China University of Petroleum (East China), Qingdao, 266580, Shandong, China*
[b] *Qingdao Institute of Marine Geology, Ministry of Natural Resources, Qingdao, 266237, Shandong, China*

A R T I C L E I N F O

Article history:
Received 20 July 2025
Received in revised form
1 December 2025
Accepted 2 February 2026
Available online 6 February 2026

Edited by Meng-Jiao Zhou

Keywords:
Full waveform inversion
Gradient decomposition
BEEMD
Migration component
Tomography component

A B S T R A C T

Full waveform inversion (FWI) gradient inherently contains both tomographic and migration components, which are responsible for updating large-scale background velocity and small-scale structural details, respectively. A key challenge in FWI is to decouple these components to ensure a robust, multi-scale inversion strategy. To overcome this, we propose a novel gradient decomposition method based on bidimensional ensemble empirical mode decomposition (BEEMD). In contrast to conventional bidimensional empirical mode decomposition (BEMD), the proposed method employs a noise-assisted ensemble averaging scheme to effectively mitigate mode-mixing artifacts. In this framework, the migration and tomographic components are respectively derived from fine-scale bidimensional intrinsic mode functions (BIMFs) and the large-scale residual. This decoupling allows the tomographic component mitigates cycle-skipping in the presence of inaccurate initial models, whereas the migration component accelerates convergence to a high-resolution solution once the background is well-defined. Sensitivity kernel analyses and numerical experiments on both synthetic and the Marmousi models demonstrate that the proposed method possesses superior stability and robustness compared to conventional approaches.

1. Introduction

Full waveform inversion (FWI) is a powerful method to construct high-resolution subsurface velocity models (Tarantola, 1984; Virieux and Operto, 2009). It achieves this by iteratively minimizing the residuals between observed and synthetic seismic wavefields (Pratt, 1999; Plessix, 2006). However, its performance is critically dependent on an accurate initial velocity model. This complexity stems from several key factors, including the model parameterization, the initial velocity field, and the quality of the seismic records. Moreover, FWI constitutes a strongly nonlinear and underdetermined inverse problem characterized by data sparsity, model parameterization errors, unknown source wavelets, and solution non-uniqueness. To overcome these challenges, the research community has developed various solution approaches (Chi et al., 2014; Métivier et al., 2018; Yong et al., 2018, 2019; Wang et al., 2019; Kazei et al., 2021; Yang et al., 2020; Li and Alkhalifah, 2020; Liu et al., 2020; Fu et al., 2023; Chen and Huang, 2024).

Mora (1989) pointed out that FWI includes two modes: migration and tomography. The migration gradient is that which updates the reflector details of the velocity model to enhance resolution, whereas the tomography gradient is that which updates the large-scale information to mitigate cycle-skipping due to a lack of low-frequency data (Yao et al., 2020). To effectively separate the migration and tomography components, current separation methodologies can be categorized into four primary approaches: (1) up/down wavefield separation; (2) scattering angle filtering; (3) reflection-waveform inversion (RWI); (4) bidimensional empirical mode decomposition (BEMD)-based FWI.

The source and receiver wavefields are first decomposed into up- and down-going components based on their propagation directions. By cross-correlating the four resulting wavefields, the tomographic component of the gradient is constructed from wavefields with the same propagation direction, while the migration component is derived from those with opposite directions. During the early stages of inversion, the tomographic component is emphasized to update the long-wavelength structures in the velocity model. This approach helps establish a favorable initial model for conventional FWI, thereby improving

* Corresponding author.
E-mail address: jphuang@upc.edu.cn (J.-P. Huang).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.02.002
1995-8226/

the final inversion accuracy (Liu et al., 2011; Chi et al., 2017; Wang et al., 2016). Fei et al. (2015) developed this imaging condition into a de-primary RTM method that performs the necessary wavefield separation without requiring the storage of complete wavefields. This innovation significantly reduces the computational memory overhead, making the approach more practical for large-scale applications. Irabor and Warner (2016) proposed splitting the wavefield into two parts (one traveling vertically and the other horizontally), and then applied the method to separate these two parts into upward and downward waves. Alkhalifah (2014) decomposed the FWI gradient using a scattering angle filter based on a criterion where a scattering angle close to 180° yields the tomography component, while other angles produce the migration component. Wu and Alkhalifah (2017) proposed a scattering angle enrichment filter to direct the migration component of the gradient to the perturbation model and the tomography component to the background model. Yao et al. (2018) proposed to Fourier transform the source and residual wavefields into the wavenumber domain and then achieve the scattering angle filtering by measuring the angle of the elements in the wavefields. Yao et al. (2019a) developed a non-stationary smoothing method with Gaussian filters based on the characteristics of the scattering angle filtering and efficiently extracted the tomography component. Reflection-waveform inversion (RWI) can also effectively separate migration and tomography components (Xu et al., 2012; Sun et al., 2017; Yao and Wu, 2017; Yao et al., 2019b; Ren et al., 2019; Huang et al., 2015). Xu et al. (2022) proposed the Gauss–Newton-based second-order optimization RWI method, and applied the method to streamer seismic data from the East China Sea.

Nie et al. (2025) recently introduced a novel approach for separating FWI gradients, utilizing the bidimensional empirical mode decomposition (BEMD) method to separate the FWI gradient into the tomography and migration component. The small-scale bidimensional intrinsic mode functions (BIMFs) are combined into the migration component, while the large-scale residual component is considered to be tomographic component. The BEMD is an adaptive, data-driven algorithm designed to decompose complex, non-stationary signals into a collection of intrinsic oscillatory components. The method extracts simple oscillatory modes called BIMFs through an iterative sifting process. The core principle requires each BIMF to meet two criteria: the numbers of local extrema and zero-crossings must be equal or differ by one and the local mean defined by its upper and lower envelopes must be zero. This sifting process works by first identifying all the local extrema of the signal, then constructing its upper and lower envelopes via interpolation. The mean of the envelopes is subtracted from the signal iteratively until the IMF criteria are met. Once this IMF is extracted, the process repeats on the residual signal until it becomes monotonic. The final outcome is a complete and adaptive decomposition of the original signal into a finite set of IMFs, arranged from the finest to the coarsest temporal scales, plus a final residue representing the overall trend (Huang et al., 1998). However, Huang et al. (1999) found there was the problem of mode mixing with empirical mode decomposition (EMD) when the data contained intermittency, and it made the intrinsic mode functions (IMFs) lose its physical meaning. To alleviate the problem of mode mixing, Wu and Huang (2009) introduced the ensemble empirical mode decomposition (EEMD) method. The key idea behind EEMD relies on averaging the modes obtained by EMD applied to several realizations of Gaussian white noise added to the original signal.

In this study, we propose an application of bidimensional ensemble empirical mode decomposition (BEEMD) as a tool for separating the tomographic and migration components in FWI. BEEMD adaptively decomposes the FWI gradient into multi-scale subcomponents through its noise-assisted ensemble framework. We constitute the migration component by combining fine-scale BIMFs and identify the residual component as the tomographic component. High-wavenumber gradient accelerates convergence and improve accuracy during high-frequency inversion, while low-wavenumber gradient stabilizes low-frequency deficient inversions by preventing convergence to local minima. Numerical experiments demonstrate that BEEMD serves as an effective method for separating these two components.

2. Theory

2.1. Basic methodology of FWI in the time domain

The acoustic wave equation can be expressed as

$$\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2}-\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial z^2}\right)=f, \tag{1}$$

where v represents velocity, u represents the wavefield, t represents time, f represents the source term, x represents distance, and z represents depth. The objective function of FWI based on the L_2 norm can be defined as follows (Tarantola, 1984):

$$\phi(m)=\frac{1}{2}\|d_{\text{obs}}-d_{\text{syn}}\|_2^2, \tag{2}$$

where m denotes the velocity model, d_{obs} is the observed data, and d_{syn} represents the synthetic data. According to the acoustic wave equation, the function for the gradient is given by

$$G(v)=\frac{\partial\phi(v)}{\partial v}=-\frac{2}{v^3}\frac{\partial^2 u}{\partial t^2}u*, \tag{3}$$

where $u*$ represents adjoint wavefield, which is computed through the backpropagation of data residuals. Then we use the conjugate gradient algorithm and get the updating process of the velocity model as

$$\begin{cases} m_{k+1}=m_k+\alpha_k g_k \\ g_k=\begin{cases} -QG_k, k=1 \\ -QG_k+\beta_k y_{k-1}, k\geq 2 \end{cases} \end{cases}, \tag{4}$$

where k is the number of iterations, α denotes the step length, g denotes the conjugate gradient, G stands the gradient, β represents the correction factor of the conjugate gradient, and Q is for the gradient preconditioning operator.

2.2. The BEEMD method

With the BEMD method, the original image is decomposed into several BIMFs and a residual component. The procedure for extracting the j-th BIMF component of the 2D signal $W(x, y)$ is as follows:

1. Initialization: where $W(x, y)$ represents the original signal, $\text{Res}_{j-1}(x, y)$ denotes residual after j-th BIMF extraction,

$$I_j(x,y)=\begin{cases} W(x,y), j=1 \\ \text{Res}_{j-1}(x,y), j\geq 2; \end{cases} \tag{5}$$

2. Extract local extrema of $I_j(x, y)$
3. Construct envelope surfaces: Upper envelope $e_{\text{upper}}(x, y)$ via interpolation of maxima and lower envelope $e_{\text{lower}}(x, y)$ via RBF interpolation of minima
4. Compute the envelope mean:

$$\overline{E}_j(x,y) = \frac{e_{\text{upper}}(x,y) + e_{\text{lower}}(x,y)}{2}; \tag{6}$$

5. Update candidate component: k represents the sifting iteration number,

$$H_j^{(k)}(x,y) = I_j(x,y) - \overline{E}_j^{(k)}(x,y); \tag{7}$$

6. Calculate the SD: If the stopping criterion for SD $\leq \sigma$ is satisfied, set $\text{BIMF}_j(x,y) = H_j^{(k)}(x,y)$ Otherwise, reset $H_j^{(k)}(x,y) = I_j(x,y)$ and iterate through steps (2)–(6) again
7. Update the residual:

$$\text{Res}_j(x,y) = \text{Res}_{j-1}(x,y) - \text{BIMF}_j(x,y). \tag{8}$$

BEMD demonstrates partial efficacy while also exhibiting persistent challenges requiring resolution. A primary limitation is the frequent presence of leakage waves, caused by two-dimensional mode mixing during the BEMD decomposition (Huang et al., 1999). This phenomenon occurs when intrinsic mode constituents improperly migrate across IMF boundaries. To overcome 2D mode mixing, BEEMD introduces controlled noise perturbation. Its fundamental principle establishes components as ensemble means derived through parallel trials. Each trial incorporates the target signal superimposed with finite-amplitude Gaussian white noise.

BEEMD achieves scale separation with enhanced precision while eliminating subjective intervention criteria. This ensemble strategy exploits the statistical properties of white noise: stochastic variations across trials cancel during sufficient ensemble averaging, yielding convergent estimates of authentic signal components. The BEEMD procedure formalizes as follows,

1. Noise injection: Supply some white noise $S(x, y)$ to 2D signal $W(x, y)$, N represents the number of trials (ensemble number),

$$I_N(x,y) = W(x,y) + S_N(x,y); \tag{9}$$

2. BEMD decomposition: Decompose via BEMD technique to obtain the first decomposed component using Eqs. (5)–(9), the interpolation method selected is radial basis function (RBF) interpolation, $\text{BIMF}_{j,N}$ represents the j-th BIMF from N ensemble trials,

$$I_N(x,y) = \sum_{j=1}^{J_N} \text{BIMF}_{j,N}(x,y) + \text{Res}_N(x,y); \tag{10}$$

3. Ensemble averaging: Average all the trial results by taking the mean of the BIMFs of the same order to obtain the BEEMD-derived BIMF components, the residual is similarly averaged, $\overline{\text{BIMF}}$ represents the averaged BIMF from N ensembles, $\overline{\text{Res}}$ represents the averaged Res from N ensembles,

$$\overline{\text{BIMF}}_j(x,y) = \frac{1}{N}\sum_{n=1}^{N} \text{BIMF}_{j,n}(x,y), j = 1, 2, ..., J, \tag{11}$$

$$\overline{\text{Res}}(x,y) = \frac{1}{N}\sum_{n=1}^{N} \text{Res}_n(x,y), \tag{12}$$

4. Signal reconstruction: The final decomposition is obtained. In BEEMD, the original signal admits the approximate reconstruction.

$$W(x,y) = \sum_{j=1}^{J} \overline{\text{BIMF}}_j(x,y) + \overline{\text{Res}}(x,y). \tag{13}$$

To better illustrate the issue of mode mixing, we performed a comparative analysis using both EMD and EEMD on a one-dimensional test signal (EMD and EEMD represent the one-dimensional versions of BEMD and BEEMD). This composite signal, shown in Fig. 1, consists of several distinct components. The signal $s(t)$ is composed of $x_1(t)$, $x_2(t)$, $x_3(t)$, and $x_4(t)$, which represent sinusoidal components at 60 Hz, 8 Hz, and 4 Hz, and Gaussian white noise, respectively. As can be seen from the area within the black box, the EMD result of signal $s(t)$ in Fig. 2 reveals inconsistent frequency content within IMF2, indicating clear mode mixing. In contrast, the EEMD result in Fig. 3 shows well-defined and consistent frequencies across all IMFs along with effective mitigation of the mode mixing phenomenon.

2.3. Gradient decomposition using BEEMD

The FWI gradient is decomposed via BEEMD into the summation of BIMFs and a residual component.

$$G = \sum_{j=1}^{J} \overline{\text{BIMF}}_j + \overline{\text{Res}}. \tag{14}$$

We interpret high-frequency components as the migration gradient and low-frequency components as the tomographic component. Specifically, the sum of the first two BIMFs is defined as the migration component, whereas the combination of all subsequent BIMFs and the residual constitutes the tomographic component. The reason for selecting the first two BIMFs to form

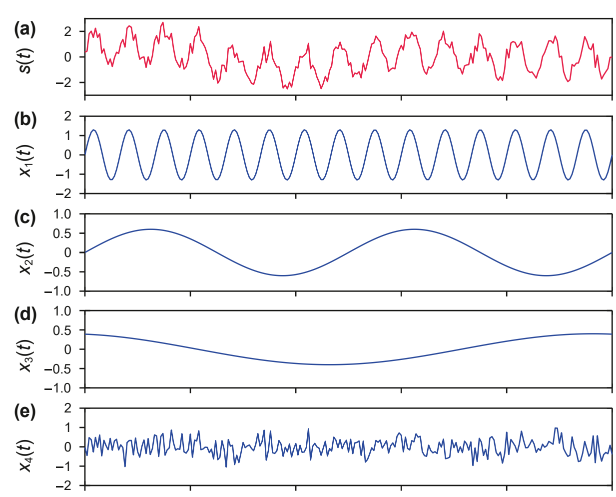

Fig. 1. **(a)** Synthetic Signal $s(t)$, **(b)** $x_1(t)$, 60 Hz sinusoidal signal, **(c)** $x_2(t)$, 8 Hz sinusoidal signal, **(d)** $x_3(t)$, 4 Hz sinusoidal signal, **(e)** $x_4(t)$, Gaussian white noise.

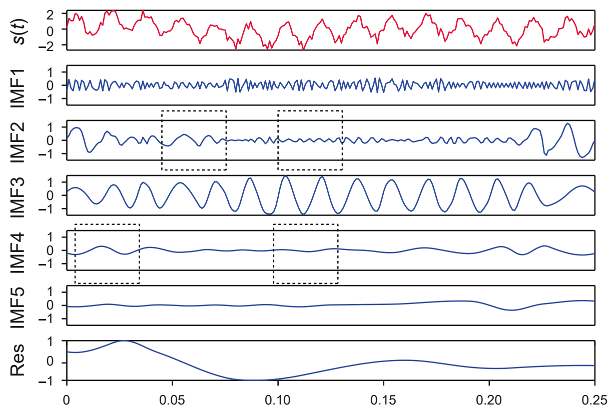

Fig. 2. IMFs and residual components derived from the EMD of $s(t)$.

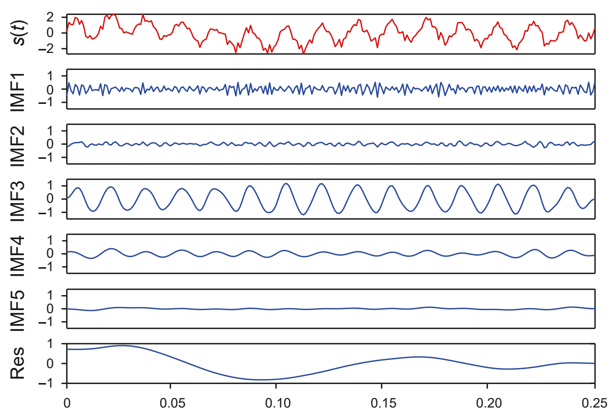

Fig. 3. IMFs and residual components derived from the EEMD of $s(t)$.

the migration component is as follows: while BIMF1 contains the highest-frequency content, it is often dominated by noise and ultra-short-wavelength scattering. In contrast, BIMF2 typically carries the dominant energy of the primary reflection signals. The combination of BIMF1 and BIMF2 ensures a more complete representation of the high-wavenumber reflective wavefield.

$$G_{\mathrm{mig}} = \sum_{j=1}^{p} \mathrm{BIMF}_j, \tag{15}$$

$$G_{\mathrm{tomo}} = \sum_{j=p+1}^{J} \mathrm{BIMF}_j + \mathrm{Res}. \tag{16}$$

where p is the number of BIMF that combine into migration component, and $p \le J$.

We summarize the workflow of FWI based on gradient BEEMD. Each iteration consists of the following steps: (1) Forward modeling to generate synthetic seismograms and compute residuals with observed data; (2) backpropagation of receiver residuals to obtain residual wavefields; (3) cross-correlation of forward and adjoint wavefields in time to compute the gradient; (4) decompose the gradient into tomography component and migration component using BEEMD; (5) update the velocity model using Eq. (4).

3. Numerical examples

In this section, we perform sensitive kernel inversion iterations and compare the results with those obtained using BEMD decomposition to highlight the advantages of the BEEMD method. Subsequently, BEEMD-FWI is applied to an anomaly model, the SEAM Arid model and the Marmousi model to validate its correctness and effectiveness. Experimental results are compared with those from conventional FWI, further confirming that BEEMD-FWI can effectively reconstruct the background velocity model and improve inversion accuracy.

3.1. Sensitivity kernel test

This test employs a simple two-layer model to validate the effectiveness of the BEEMD method for gradient decomposition. The model dimensions are 200 by 200 grid points with a sampling interval of 10 m. In the true model, the P-wave velocities of the upper and lower layers are 2.0 and 3.0 km/s, respectively. The reflection interface is positioned at a depth of 1.5 km. The initial model shares identical structural properties with the true model, except that the lower layer velocity is set to 2.95 km/s. A Ricker wavelet with a dominant frequency of 10 Hz serves as the seismic source, located at coordinates (0.5 km, 1.0 km), while the receiver is positioned at (1.5 km, 1.0 km). Fig. 4 displays the sensitivity kernel of reflected waves in conventional FWI. Red and blue dots denote the source and receiver positions, respectively. The “rabbit-ear” shaped tomographic component is formed by contributions from source-interface-receiver reflection paths, whereas the migration component manifests as a semi-circular pattern. Conventional FWI is significantly influenced by the migration component, impeding effective inversion of the background velocity model. Therefore, enhancing the tomographic functionality of FWI requires separately extracting the tomographic component.

Fig. 5(a)–(c) represent two BIMF components of different scales and the residual component obtained using BEMD. These BIMFs reflect information about different temporal scales and frequency components within the original signal. Typically, the BIMF1 represents the highest frequencies in the signal, with subsequent BIMFs exhibiting progressively lower frequencies. The ideal decomposition result would isolate each distinct frequency component into a single BIMF. In practical applications, a single BIMF frequently contains multiple frequency components, or a

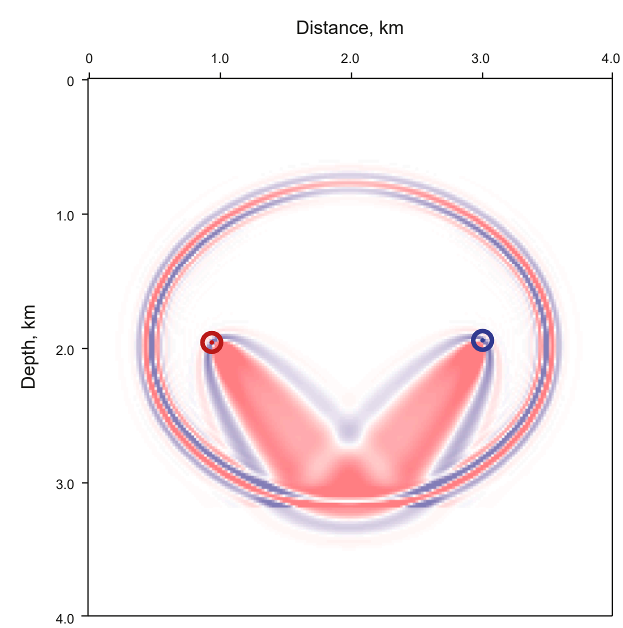

Fig. 4. The sensitivity kernel in conventional FWI.

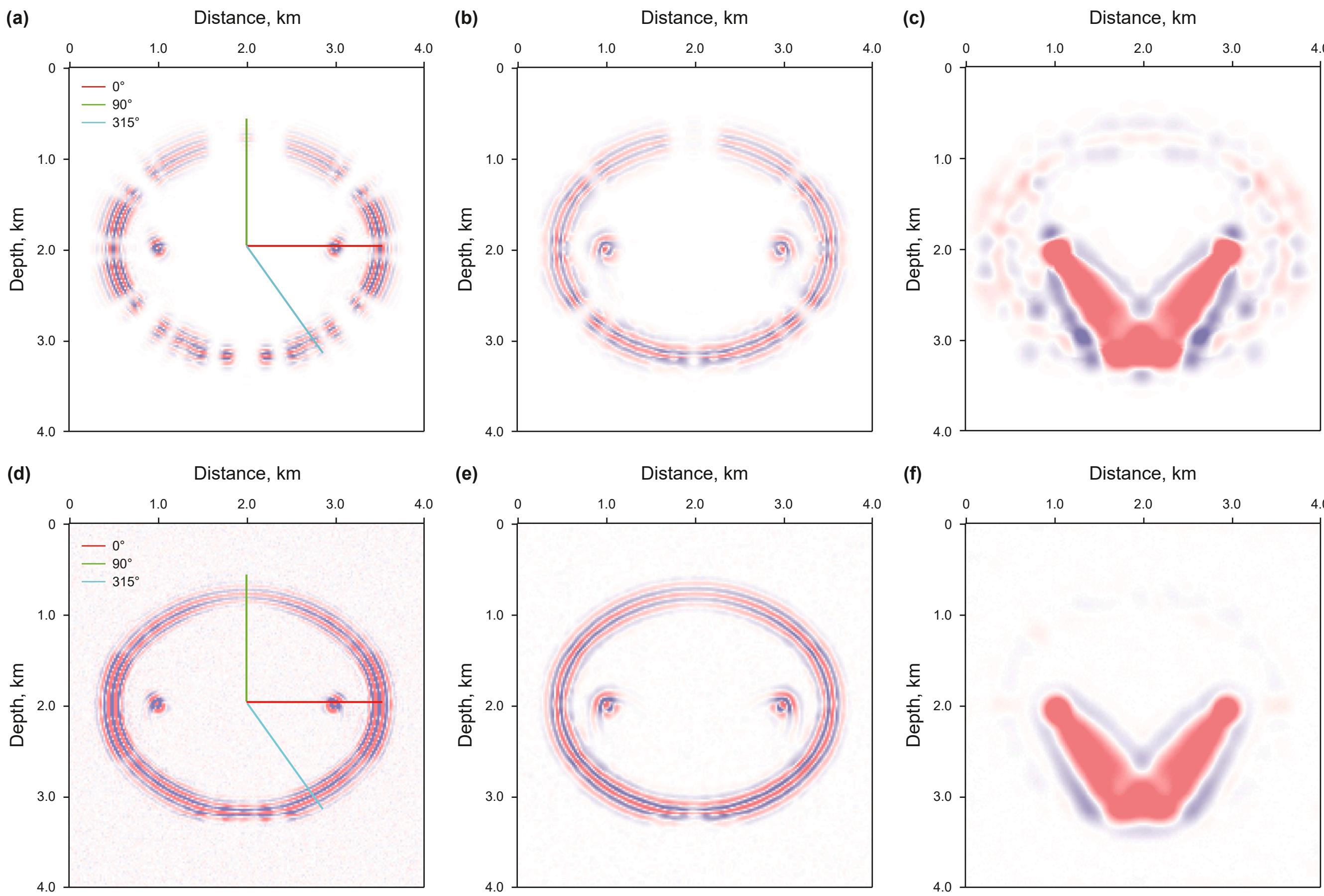

Fig. 5. The sensitivity kernels decomposed by (a)–(c) BEMD and (d)–(f) BEEMD. **(a)** BEMD- BIMF1, **(b)** BEMD-BIMF2, **(c)** BEMD-residue, **(d)** BEEMD-BIMF1, **(e)** BEEMD-BIMF2 and **(f)** BEEMD-residue.

single frequency component may be distributed across several BIMFs. This issue is termed mode mixing. Mode mixing is particularly prone to occur when the frequencies of the signal components are closely spaced. The BIMF components represent small-scale information, while the residual component contains low-frequency information and possesses the largest scale. The migration component displays an approximately circular distribution in the spatial domain. Consequently, in the spectral domain, the amplitude distribution remains largely consistent across different polar angles for a given polar radius. Therefore, we selected to extract spectra along different angles in polar coordinate form. The polar angle distribution of the extracted migration component is shown in Fig. 5(a), and the obtained spectra are shown in Fig. 6(a).

In contrast, Fig. 5(d)–(f) show BIMF1, BIMF2, and the residual component obtained using BEEMD. Fig. 6 displays the polar

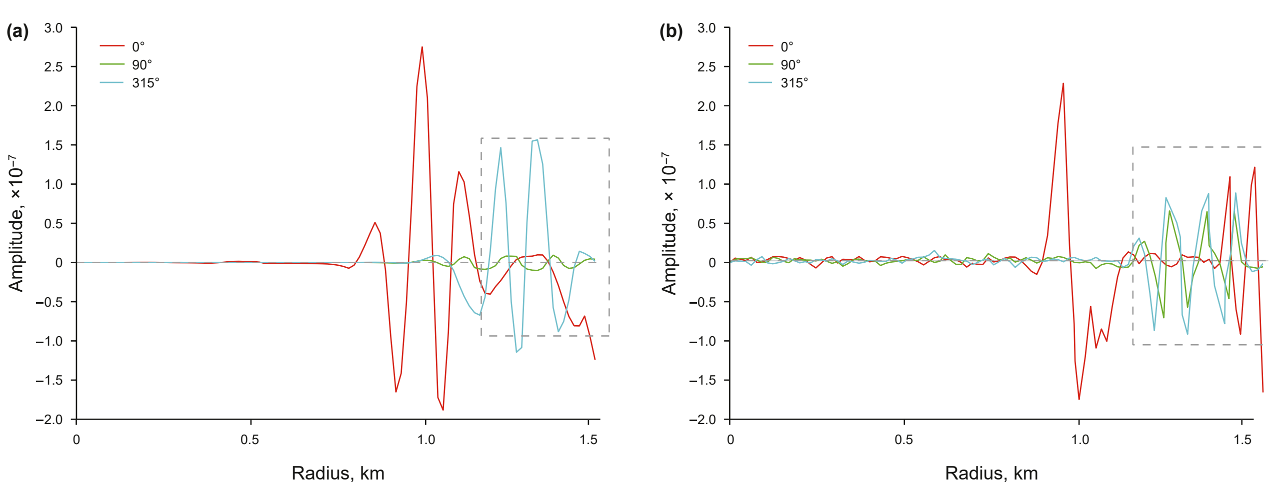

Fig. 6. The spectra in Fig. 5(a) and (d).

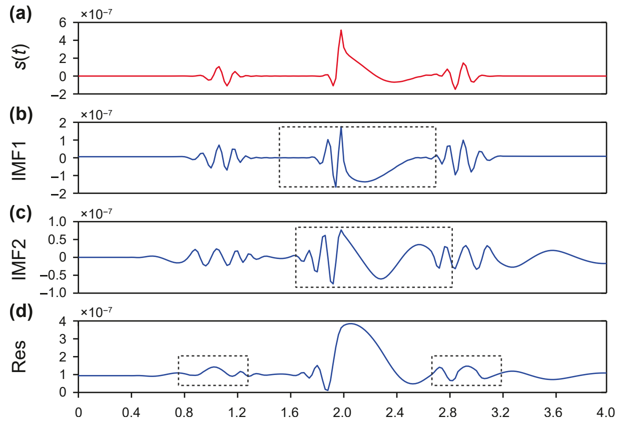

Fig. 7. IMFs and residual from EMD decomposition of the trace at x = 1.0 km.

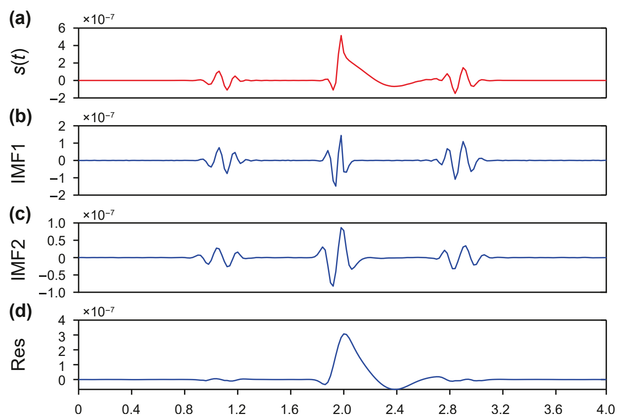

Fig. 8. IMFs and residual from EEMD decomposition of the trace at x = 1.0 km.

coordinate spectra extracted from Fig. 5(a) and (d), respectively. Since the frequency components at a polar angle of 0° (indicated by the red line) originate not only from the circular region but also include high-frequency information from the rabbit-ear areas, the analysis is confined to the area within the dashed box to enable accurate comparison of frequency variations across polar angles. As shown in Fig. 6(a), the frequency components extracted from Fig. 5(a) exhibit significant differences. In contrast, those corresponding to Fig. 5(d) in Fig. 6(b) remain largely consistent. Fig. 7 presents the decomposition results of a trace extracted at x = 1.0 km from the original gradient (Fig. 4) using the BEMD method. Fig. 8 shows the corresponding results using the BEEMD method. As can be observed in Fig. 7, the IMFs obtained via EMD exhibit inconsistent frequency content. Moreover, the residual component (the black rectangular region) retains noticeable high-frequency information, indicating the presence of mode mixing in the decomposition process. In contrast, the EEMD results in Fig. 8 demonstrate effective mode separation. All extracted IMFs exhibit consistent frequency characteristics, and the residual component is free of high-frequency remnants. This confirms the efficacy of the proposed method in mitigating mode mixing.

Consequently, all BIMFs are combined to form the migration component, while the residual component is assigned as the tomographic component. The final combined result is presented in Fig. 9, showing that BEEMD successfully separated the two components. This result effectively mitigates mode mixing. A very weak leakage of the tomographic component persists within the migration component, as indicated by the red arrow. This occurs because not all frequency components in the tomographic gradient are low in energy. Specifically, its high-frequency portion exhibits spectral characteristics that closely match those of the decomposed BIMF2, which leads to a trace residual of the tomographic signal in the migration result. Fig. 10(a)–(d) present the migration and tomographic components obtained using BEMD decomposition alongside those derived from BEEMD. The wavenumber-domain results reveal the limitations of the conventional BEMD method in gradient decomposition. Although it provides a preliminary separation of low- and high-wavenumber components, the process is incomplete and leads to significant energy leakage: low-wavenumber background information remains as artifacts in the migration gradient (Fig. 10(a)), while high-wavenumber structural details leak into the tomographic gradient (Fig. 10(b)). In contrast, BEEMD demonstrates superior scale

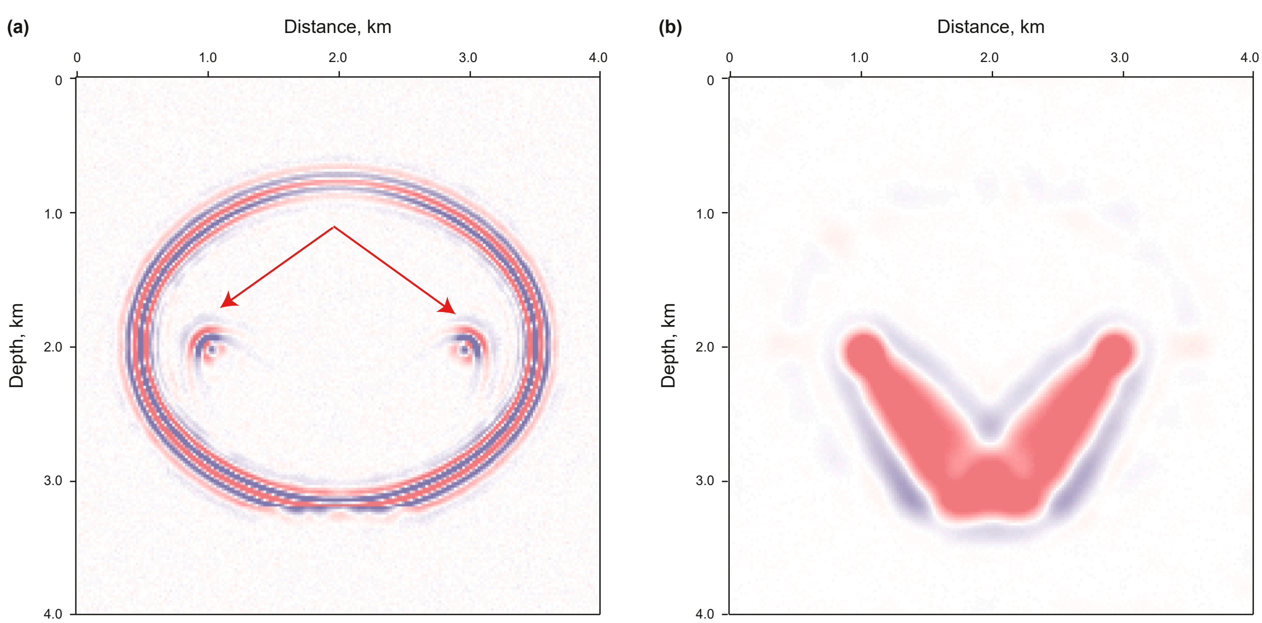

Fig. 9. **(a)** The migration component (BIMF1+BIMF2) and **(b)** the tomographic component (Residue).

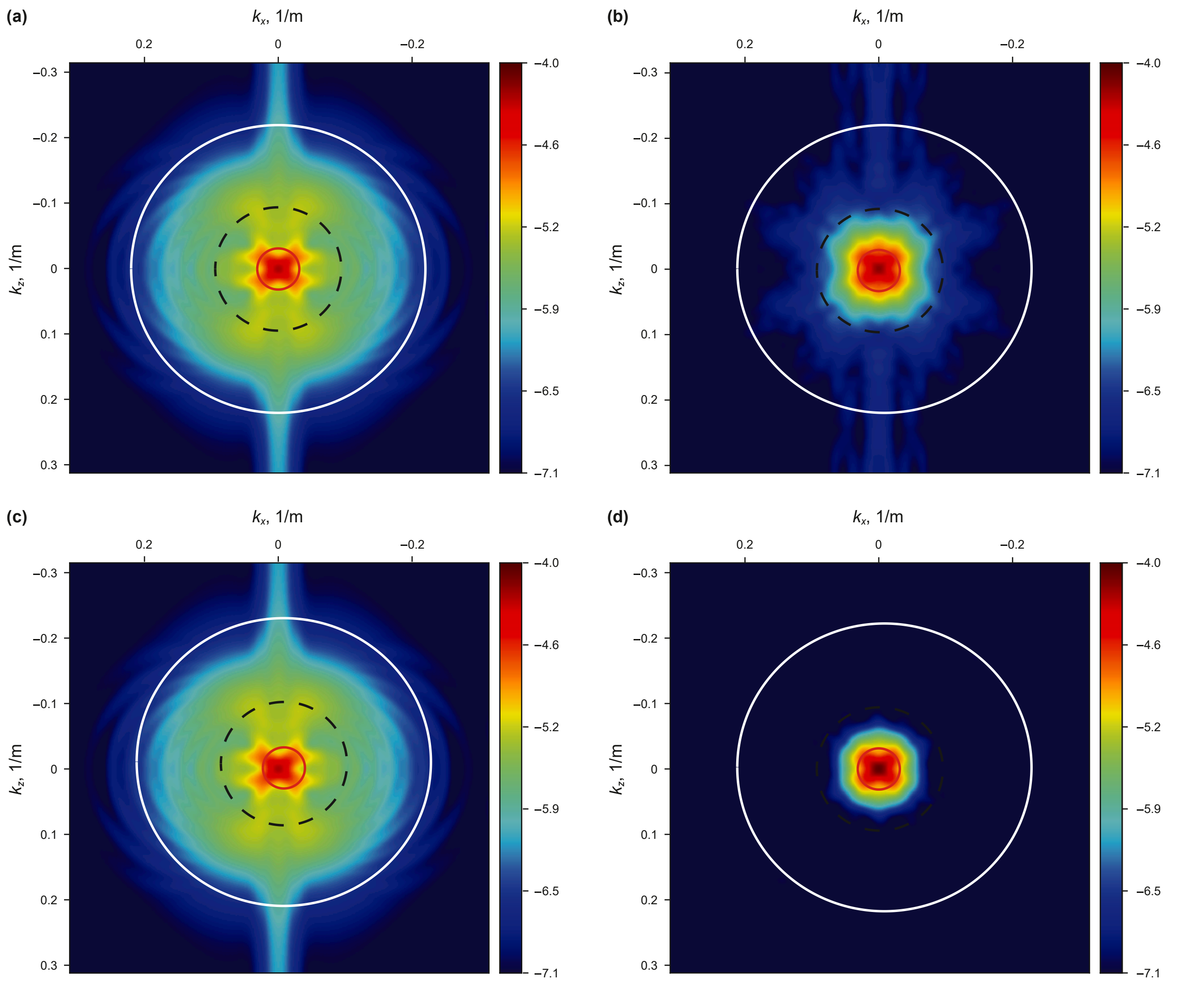

Fig. 10. The wavenumber spectra of the **(a)** migration gradient obtained using BEMD decomposition, **(b)** tomographic gradient obtained using BEMD decomposition, **(c)** migration gradient obtained using BEEMD decomposition, **(d)** tomographic gradient obtained using BEEMD decomposition.

separation capability. Its resulting migration and tomographic gradients are substantially purer, with minimal cross-contamination, thereby achieving a nearly ideal signal decoupling (Fig. 10(c) and (d)).

The advantage of the BEEMD method lies in its mitigation of mode mixing through the addition of multiple sets of independent white noise. The noise is cancelled out via ensemble averaging, preserving the effective signal. Each BIMF obtained through this decomposition represents distinct scale information from the original gradient. Consequently, the weaker correlation between different scales directly reduces leakage in the decomposition results.

3.2. Anomaly model

In this test, the proposed method is applied to a simple two-layer model. The true model (Fig. 11(a)) consists of 100 vertical and 200 horizontal grid points, with velocities of 2.5 km/s for the top layer, 2.0 km/s for the bottom layer, and a circular anomaly velocity of 2.3 km/s. The source wavelet represents a 10 Hz Ricker wavelet, with a grid spacing of 10 m. The total recording time is 3 s with a sampling of 0.015 s. The acquisition configuration comprises 21 shots and 201 receivers. The shot interval is 0.1 km, and the receiver interval is 0.01 km. The initial velocity model (Fig. 11(b)) without the circular anomaly is used for the FWI.

The results obtained from FWI using the conventional gradient and those from FWI using the tomographic component gradient based on the gradient BEEMD method after 20 iterations are shown in Fig. 12, respectively. It can be observed from the figures that the result using conventional FWI only captures detailed structures of the anomalous body but fails to resolve its internal features effectively. In contrast, inversion using the tomographic component gradient from BEEMD decomposition successfully recovers the high-velocity anomalous body and provides a favorable initial velocity model for subsequent inversion. Fig. 13 presents the conventional FWI results after 40 iterations using the velocity models from Fig. 12(a) and (b) respectively as initial models. The results demonstrate that the anomaly model inverted using the BEEMD gradient method is significantly closer to the true model and exhibits a more distinctive shape definition compared to the result from conventional FWI.

To clearly demonstrate the differences between these two methods, Fig. 14 presents a comparison of velocity profiles extracted at the same location from the true velocity model.

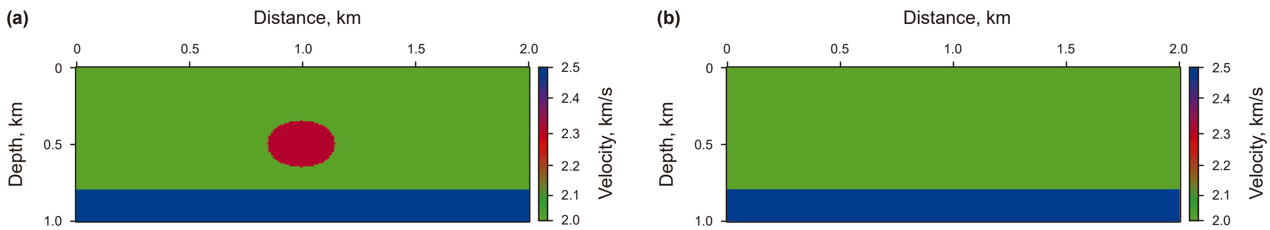

Fig. 11. Anomaly model. **(a)** True model with a high-velocity anomaly and **(b)** initial model.

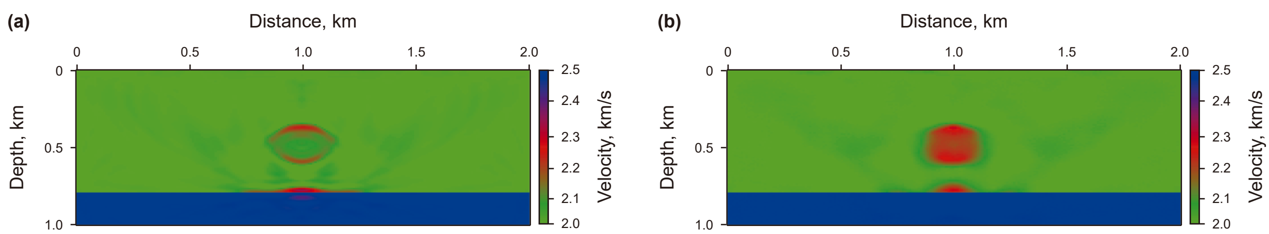

Fig. 12. Inversion results after 20 iterations using **(a)** conventional FWI and **(b)** BEEMD-FWI.

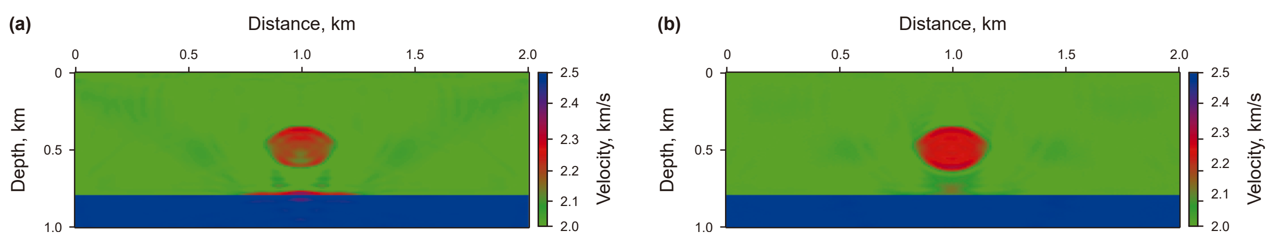

Fig. 13. FWI results after40 iterations. The initial models of **(a)** and **(b)** are shown in Fig. 12(a) and (b), respectively.

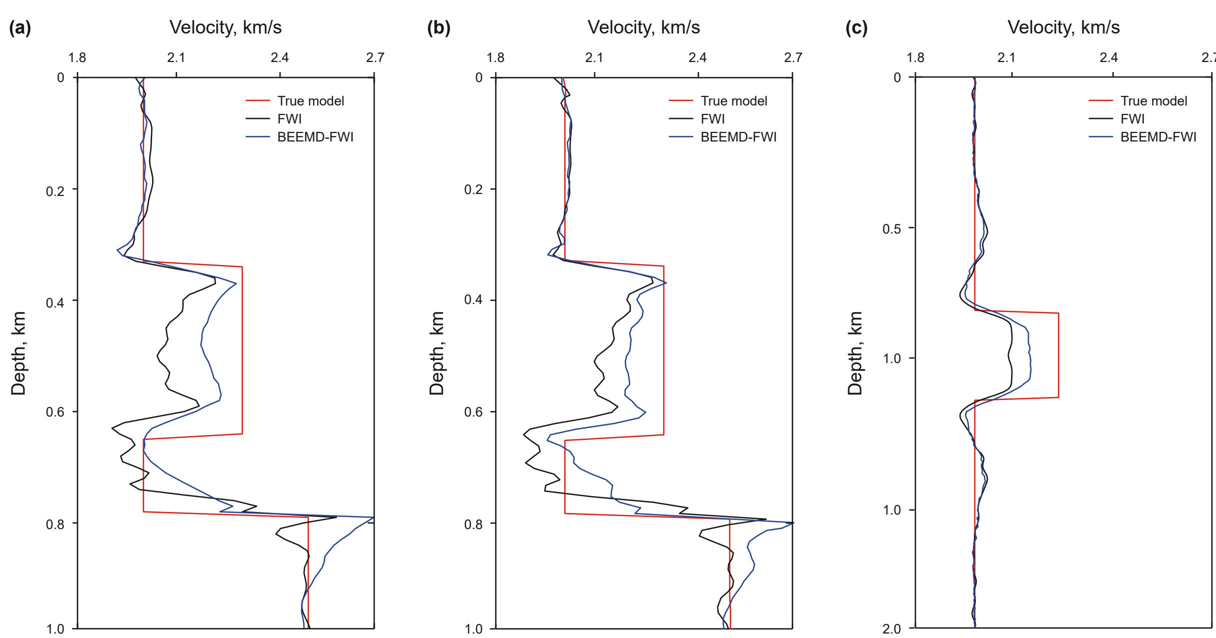

Fig. 14. Comparison of the velocity curves. These traces are extracted from **(a)** Figs. 11(a), 12(a) and (b) at a distance of 1.0 km, **(b)** Figs. 11(a), 13(a) and (b) at a distance of 1.0 km, and **(c)** Figs. 11(a), 13(a) and (b) at a depth of 0.5 km.

Fig. 14(a) extracts traces located at a horizontal distance of 1.0 km from the velocity model. Fig. 14(b) extracts traces located at a horizontal distance of 1.0 km from the velocity model. Fig. 14(c) extracts traces located at a depth of 0.5 km from the velocity model. The red line represents the true velocity model, the black line represents the result obtained from FWI using the conventional gradient, and the blue line represents the result obtained from FWI using the tomographic component gradient derived from BEEMD decomposition. It can be seen that at the location of the anomalous body, the blue line is closer to the red line than the black line. Therefore, using the tomographic component gradient from BEEMD-FWI leads to a more successful reconstruction of the high-velocity anomalous body. This result provides a favorable initial velocity model for the subsequent inversion.

3.3. Seam arid model

We employ a 2D slice extracted from the SEAM Arid model to test the effectiveness of the proposed method (Oristaglio, 2015). This model presents a notoriously demanding scenario for FWI, particularly within the highly heterogeneous near-surface region. The complex scattering phenomena generated by these shallow, small-scale structures pose significant obstacles to achieving robust FWI convergence.

For this test, we employ an array of 20 sources and 400 receivers located uniformly at a depth of 10 m. A 10 Hz Ricker wavelet is used as the source. The numerical grid uses a 10 m spacing, and data are recorded for 3 s with a 1 ms time step. Fig. 15 displays the true velocity model and the smoothed initial velocity model. Based on this initial model, we performed the first iteration using conventional FWI, BEMD-FWI, and BEEMD-FWI, respectively. The gradient results are shown in Fig. 16.

As observed in Fig. 16, the gradient obtained from BEMD-FWI exhibits significant mode mixing. The gradient is contaminated with substantial high-frequency noise and spurious artifacts, which severely interfere with the extraction of the effective gradient. In contrast, BEEMD-FWI effectively mitigates the mode mixing effect by incorporating noise-assisted analysis and ensemble averaging strategies. It overcomes the scale separation errors caused by signal intermittency. Consequently, BEEMD-FWI successfully isolates and preserves the data's low-frequency components and thereby facilitates the construction of a high-quality background velocity model.

Fig. 17 illustrates the inversion results after 30 iterations for both conventional FWI and BEEMD-FWI. BEEMD-FWI updates the model by constructing tomographic gradient derived from the decomposed residuals. A comparison between Fig. 17 reveals that while conventional FWI (Fig. 17(a)) recovers the general structural outline, it suffers from significant artifact interference in the mid-to-deep zones. In contrast, the BEEMD-FWI result (Fig. 17(b)) not only substantially mitigates these artifacts but also successfully reconstructs the macro-velocity background structure of the model. BEEMD-FWI yields a smoother and more accurate deep background velocity trend with geometric morphology of the velocity interfaces closely aligning to the true model. This high-quality background model provides a robust initial velocity for

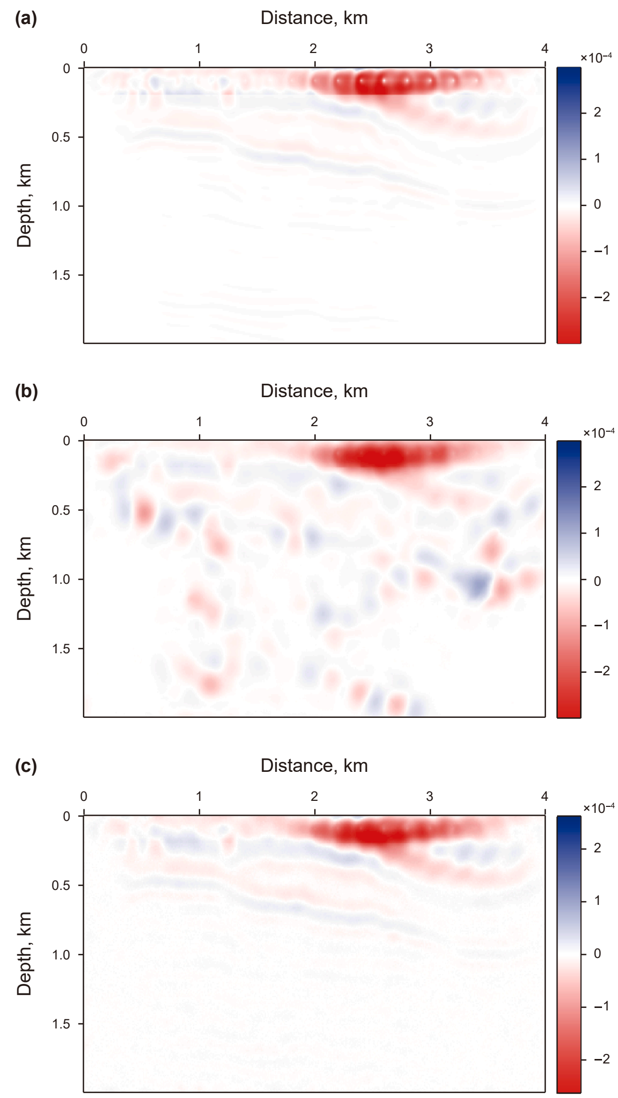

Fig. 16. The gradient after one iteration **(a)** using conventional FWI, **(b)** using BEMD-FWI and **(c)** using BEEMD-FWI.

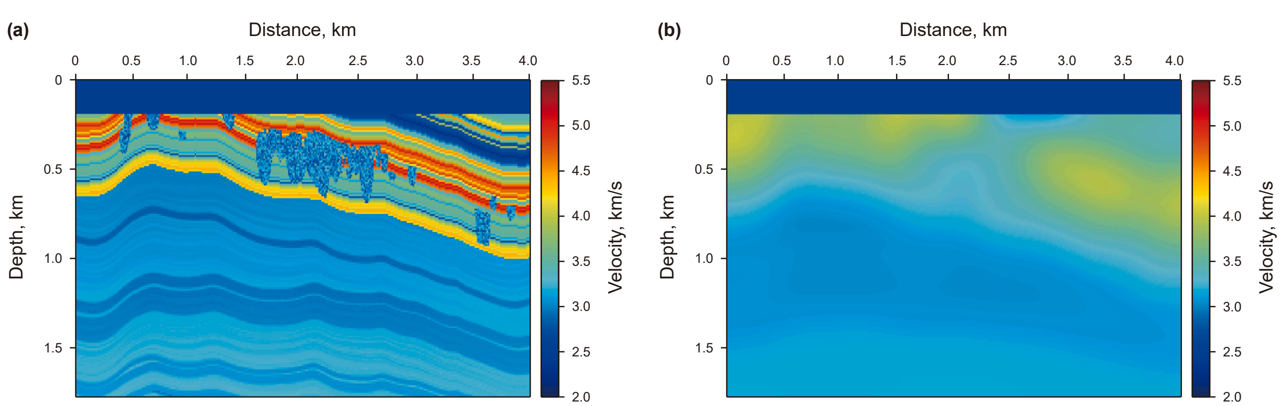

Fig. 15. Seam arid model. **(a)** The true velocity model and **(b)** the initial velocity model.

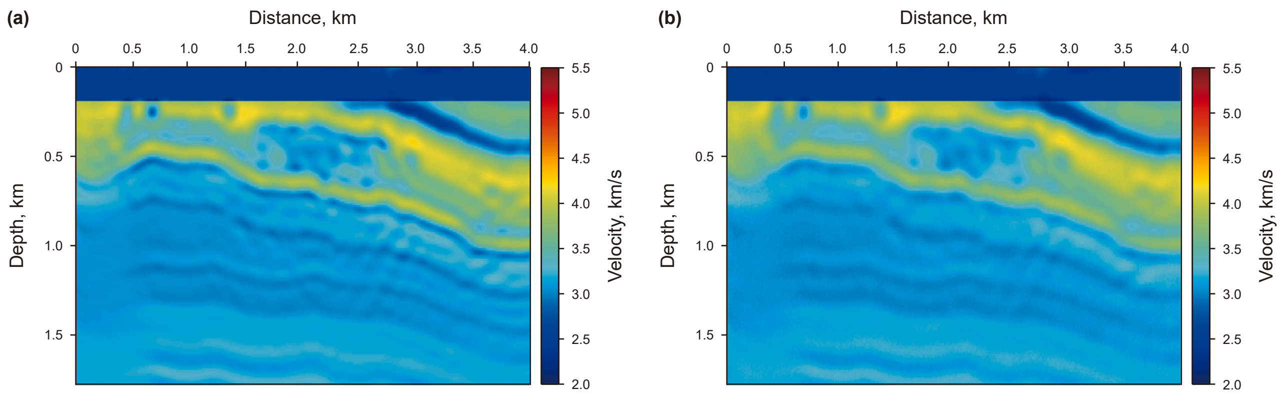

Fig. 17. Inversion results using **(a)** conventional FWI after 30 iterations, **(b)** BEEMD- FWI after 30 iterations.

subsequent inversion and effectively mitigates the risk of cycle-skipping.

To further enhance resolution based on this accurate background model we performed an additional 70 iterations. The final results are shown in Fig. 18. Fig. 18(a) represents the result of continuing with 70 iterations of conventional FWI using Fig. 17(a) as the initial model. Based on Figs. 17(b) and 18(b) was obtained by performing 50 iterations of conventional gradient updates followed by 20 iterations of migration gradient updates to enhance high-wavenumber information. As evidenced by the figures, the final accuracy achieved via the BEEMD-FWI strategy significantly outperforms the conventional method, particularly in the shallow subsurface. The conventional FWI result retains considerable artifacts in the shallow layers and exhibits blurred layer boundaries. On the other hand, the BEEMD-FWI result clearly delineates fine shallow structures and presents sharp velocity interfaces with excellent mitigation of near-surface inversion noise. This demonstrates the superior performance of the proposed method in recovering high-wavenumber details under the constraint of a large-scale background.

3.4. Marmousi model

This section employs Marmousi model to test the advantages of using the BEEMD-FWI. The true velocity model and initial velocity model are shown in Fig. 19(a) and (b), respectively. The model dimensions are 369 by 178, with a grid spacing of 10 m. The source is a 10 Hz Ricker wavelet. A total of 36 shots are deployed at a

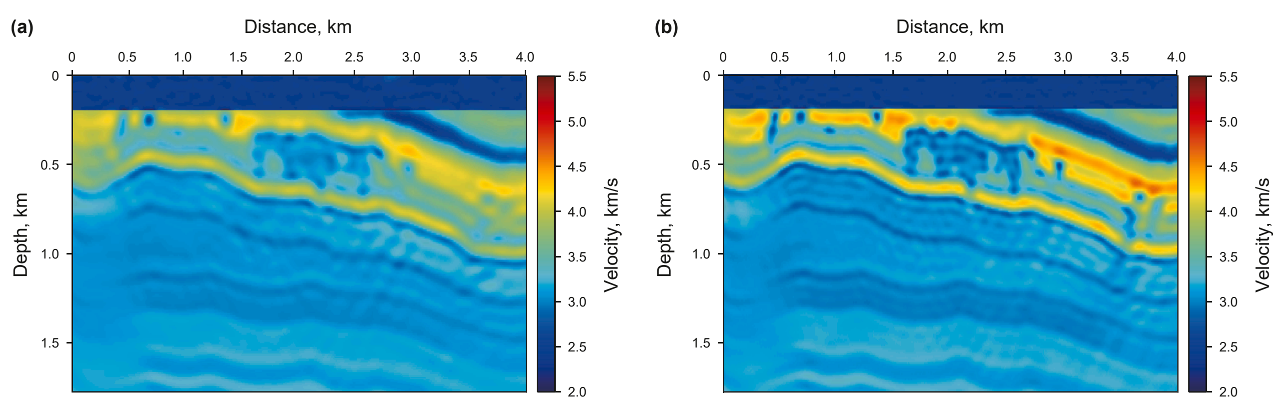

Fig. 18. FWI results after 70 iterations. The initial models of **(a)** and **(b)** are shown in Fig. 17(a) and (b), respectively.

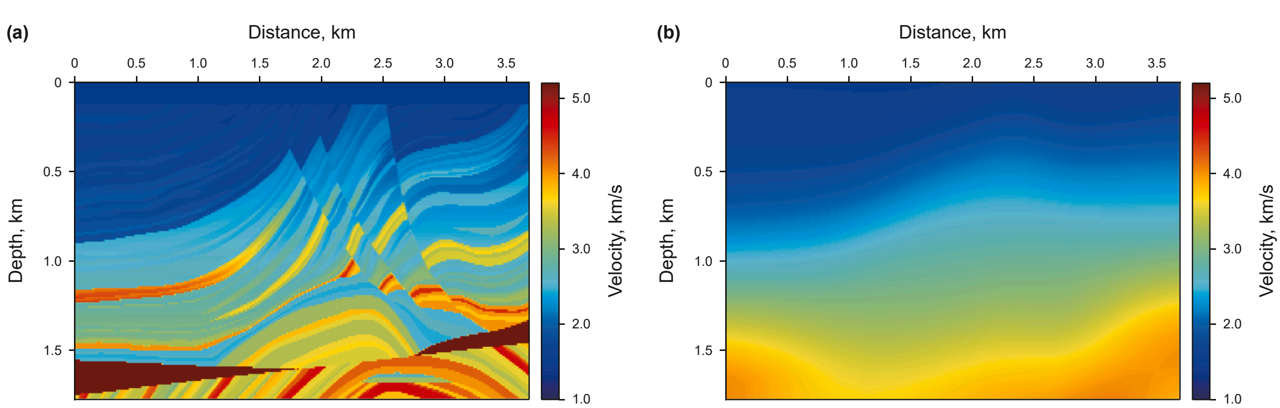

Fig. 19. Marmousi model. **(a)** The true velocity model and **(b)** the initial velocity model.

depth of 0.01 km with an interval of 0.1 km. Seismic recording employs a fixed-spread geometry with 369 receivers spaced at 0.01 km intervals. The data are recorded for 3 s with a time sampling interval of 0.001 s.

For the Marmousi model, two sets of tests are designed to comprehensively evaluate the advantages of the proposed method: Ideal Tests and Low-Frequency Missing Data Tests. The Ideal Tests use the migration gradient obtained from BEEMD decomposition for FWI, while the Low-Frequency Missing Data Tests employ the tomographic component gradient obtained from BEEMD decomposition for FWI to provide a favorable initial velocity model for subsequent inversion.

3.4.1. *Ideal tests*

Because the data contain sufficient low-frequency components, we first apply conventional FWI. Subsequently, the results obtained using this as the initial velocity model for further conventional FWI are compared with those obtained using FWI based on the migration gradient derived from BEEMD decomposition. Fig. 20(a) shows the result after 30 iterations of conventional FWI using the conventional gradient. The velocity structure in the shallow layers is largely recovered, but the velocity structure in the mid-to-deep layers is not successfully inverted. Two separate inversions are performed from this initial model, with one using the conventional gradient and the other using the migration gradient obtained from BEEMD decomposition. Fig. 20(b) and (c) present the FWI results after 120 iterations using the conventional gradient and the migration gradient derived from BEEMD decomposition, respectively. From the gray dashed lines in the figure, it can be observed that compared to the result from conventional FWI, the inversion result obtained using the migration gradient method based on BEEMD decomposition is more effective in recovering the background velocity structures within the mid-to-deep layers. This is because the migration gradient can leverage high-wavenumber components for inversion, resolving the detailed structural features of the model.

A further comparison of the inversion effectiveness can be made by extracting trace profiles from the inversion results shown in Fig. 21. These traces are extracted at distances of 0.5, 2.0, and 3.2 km. The black curve represents the inversion result obtained using the conventional gradient method, while the blue curve represents the result obtained using the migration gradient method based on BEEMD decomposition. It can be observed that within the depth range of 0–0.5 km, the velocity curves from both the conventional inversion method (black curve) and the BEEMD gradient method closely coincide with the true model (red curve). With increasing depth, the result from the BEEMD gradient method (blue line) aligns more closely with the true velocity model than that from conventional FWI (black line), due to its effective use of high-wavenumber components. Comparative analysis indicates that the proposed method achieves closer alignment with the true velocity model than conventional FWI.

As shown in Table 1, the computational cost per iteration for BEEMD-FWI is approximately three times that of conventional FWI. However, the proposed method achieves comparable accuracy in fewer iterations. This reduction in the total number of iterations effectively compensates for the higher per-iteration cost, resulting in competitive overall computational efficiency for BEEMD-FWI.

3.4.2. *Low-frequency missing data tests*

Low-frequency data is crucial for mitigating the non-linearity of the FWI. The lack of low-frequency components in seismic data often traps FWI in local minima. However, acquiring effective low-frequency information is often challenging in practical acquisition. Researching FWI methods under low-frequency-deficient conditions has therefore become an important direction. To avoid cycle-skipping phenomena caused by insufficient low-frequency components, we employ BEEMD to obtain the tomographic component gradient. This gradient effectively leverages low-wavenumber components to reconstruct large-scale structures in order to provide a favorable initial velocity model for subsequent inversion.

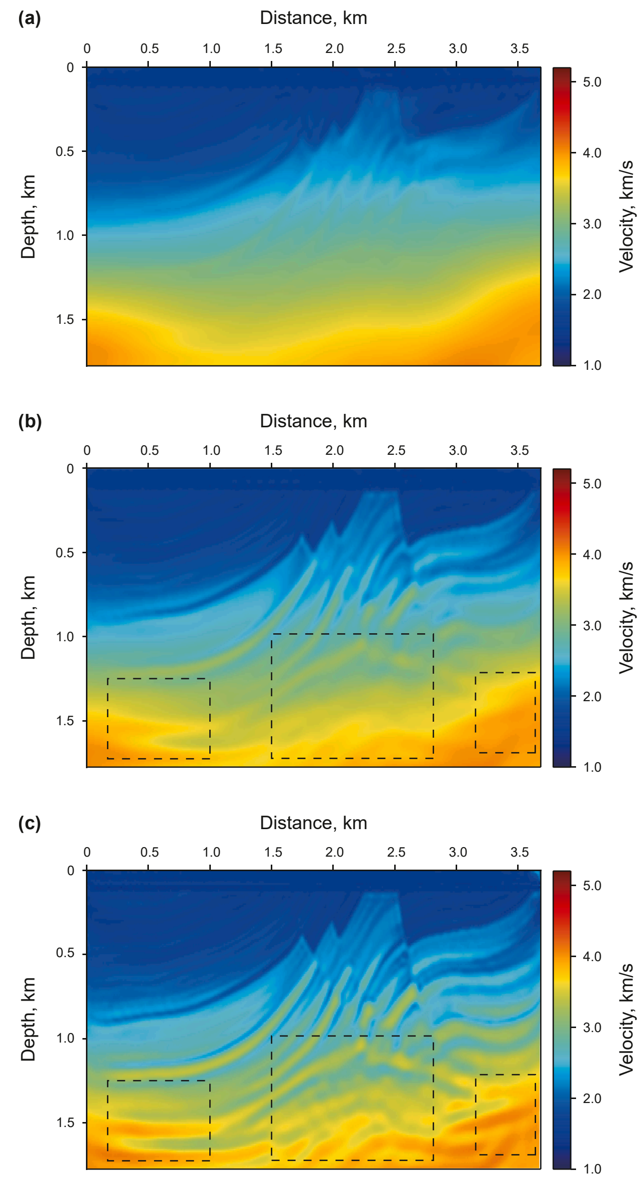

Fig. 20. Inversion results obtained using Fig. 20(a) as the initial model: **(a)** conventional FWI after 30 iterations, **(b)** conventional FWI after 120 iterations, and **(c)** BEEMD-FWI after 120 iterations.

In this test, we process the original shot record (Fig. 22(a)) by completely removing all frequency components below 2 Hz and attenuating components between 2 and 10 Hz. The resulting low-frequency-deficient shot record is shown in Fig. 22(b), with its spectrum presented in Fig. 23. The corresponding gradients after the first iteration (Fig. 24) reveal that the conventional gradient contains both tomographic and migration components, which disrupt the update of the background velocity. Fig. 25(a) and (b) display the wavenumber spectra of the gradients from Fig. 24(a) and (b). In contrast, the BEEMD-derived tomographic component

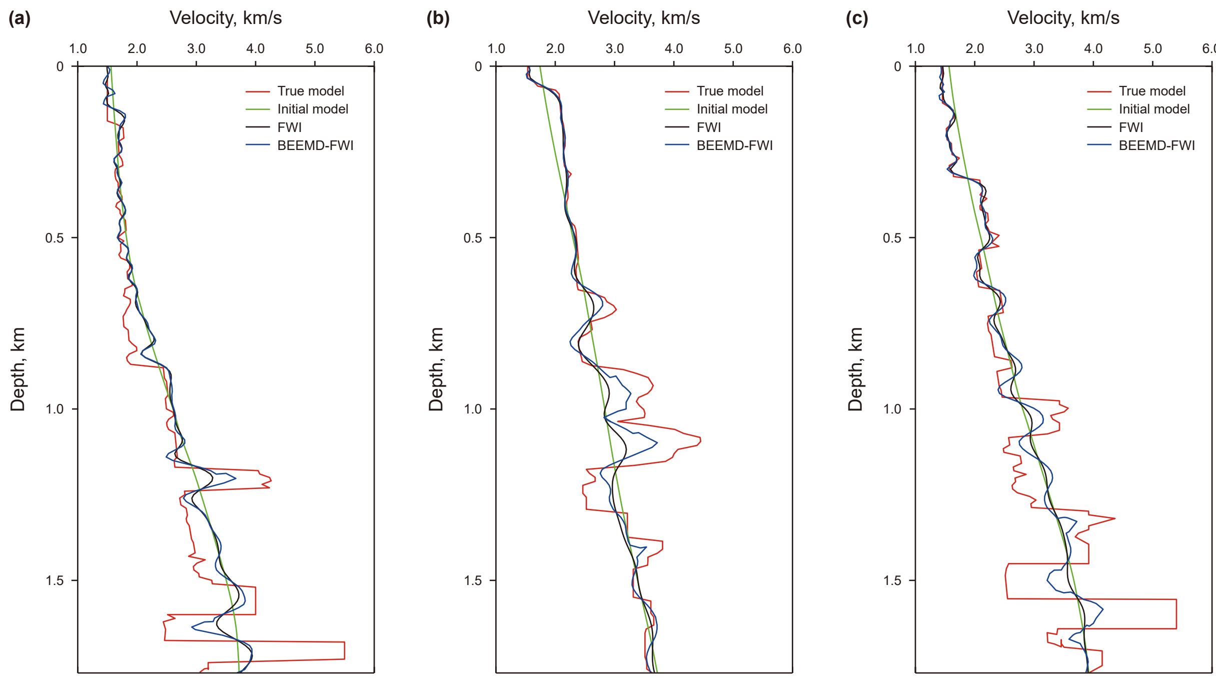

Fig. 21. Comparison of the velocity curves. These traces are extracted from the true velocity model, the initial velocity model, the conventional FWI result, and the BEEMD-FWI result at a distance of **(a)** 0.5 km, **(b)** 2.0 km, and **(c)** 3.2 km.

Table 1
Efficiency comparison per iteration between FWI and BEEMD-FWI.

Method	σ (SD)	Number of decomposed BIMFs	N (ensemble number)	Computing time of one iteration, s	Percentage of computing time, %
FWI	—	—	—	786	—
BEEMD-FWI	0.2	2	50	2403	305

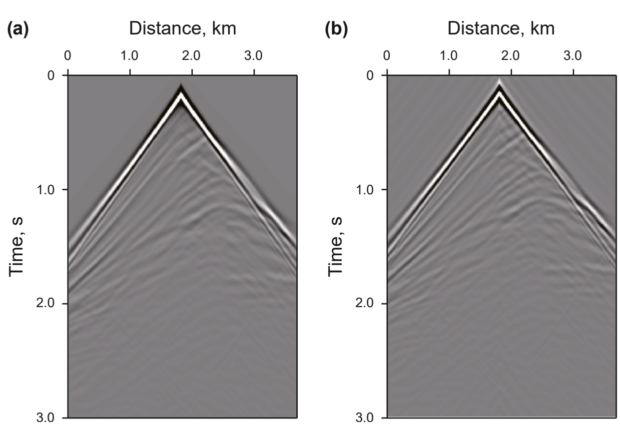

Fig. 22. Observed data. **(a)** An ideal shot record and **(b)** a low-frequency missing shot record.

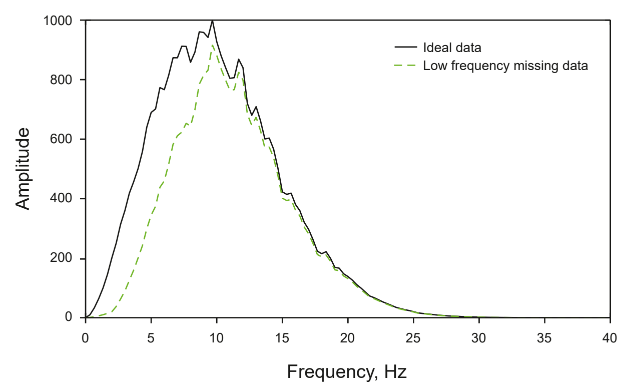

Fig. 23. Frequency spectrum for the ideal data and low-frequency missing data.

gradient exhibits enhanced smoothness and effective utilization of low-wavenumber information.

The low-frequency missing shot record is then used as the input for inversion iterations. Fig. 26(a) and (b) display the inversion results after 30 iterations using the conventional gradient and the tomographic component gradient, respectively. It can be observed that the tomographic component gradient successfully recovers the background velocity structures, establishing a superior initial model for subsequent inversion. Fig. 27(a) shows the result obtained by performing 100 iterations of conventional FWI using Fig. 26(a) as the initial velocity model. Fig. 27(b) displays the result obtained using Fig. 26(a) as the initial model, consisting of 70 iterations of conventional FWI followed by 30 iterations using migration-based gradients. The figures demonstrate that the BEEMD method yields significant improvement over conventional FWI especially in the mid-to-deep sections. The absence of low-frequency components traps conventional FWI in local minima causing its inversion result to show negligible difference from the initial velocity model. Conversely, the result obtained using FWI with the tomographic component gradient derived from BEEMD decomposition successfully recovers the velocity structure of the Marmousi model in the mid-to-deep layers.

Fig. 28 presents the normalized misfit function curves for both inversion methods. The black curve represents the conventional gradient FWI method, and the blue curve represents the FWI

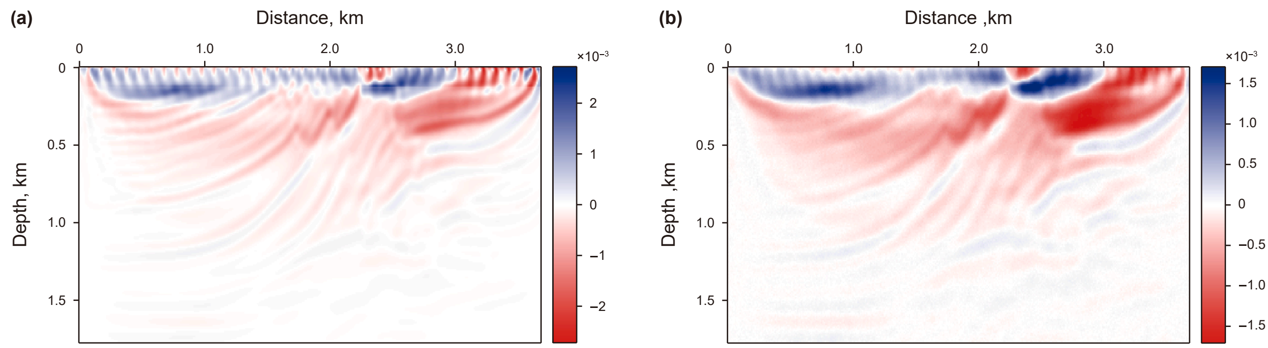

Fig. 24. The gradient after one iteration **(a)** using conventional FWI and **(b)** using BEEMD-FWI.

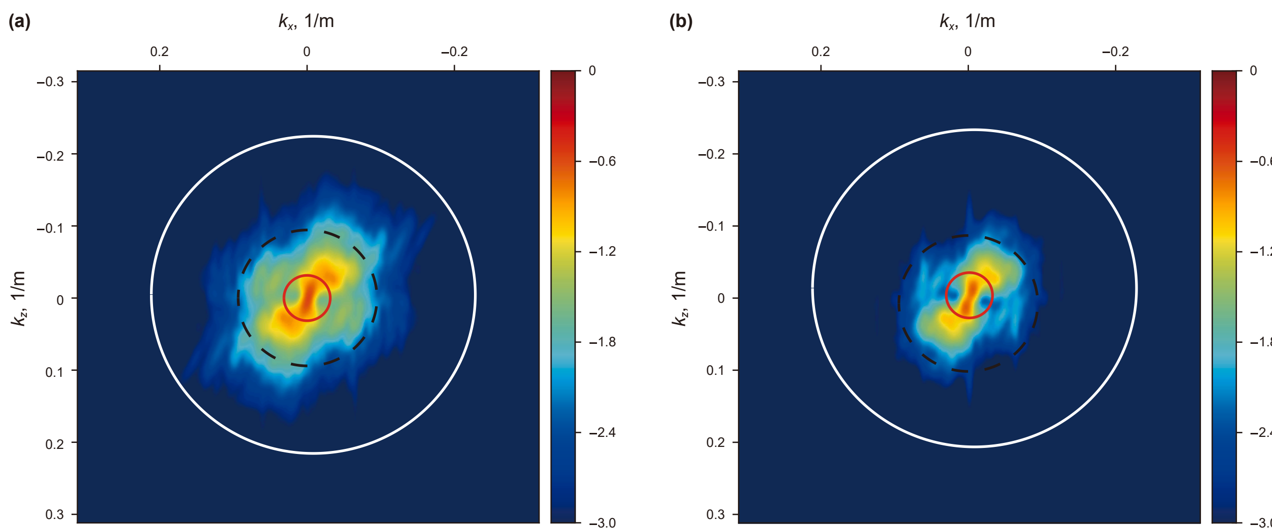

Fig. 25. The wavenumber spectra of the gradients in Fig. 24(a) and (b), respectively.

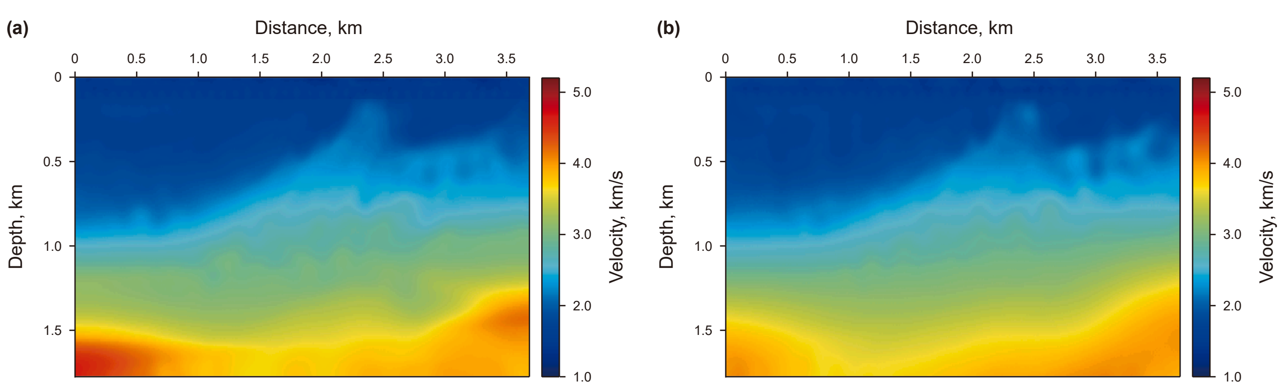

Fig. 26. Inversion results after 30 iterations for the low-frequency missing data using the **(a)** conventional FWI method and **(b)** BEEMD-FWI method.

method using the tomographic component gradient obtained via BEEMD decomposition. The circle marks the transition point after 30 iterations, where the inversion updates switch to conventional FWI using the respective results as the new initial velocity model. Within the first 30 iterations, the convergence speed of the conventional gradient method is noticeably faster than that using the BEEMD-derived tomographic component gradient. However, the convergence stagnates around iteration 30. Using the inversion results from these methods as initial models for the subsequent 120 iterations of conventional FWI, the blue curve converges more stably and rapidly than the black curve. Fig. 29 shows velocity profiles extracted at horizontal distances of 1.0, 2.0, and 3.0 km from both inversion results. The profiles employ a color code where the true velocity model is red, the initial model is green, conventional FWI is black, and the BEEMD method is blue. It can be observed that the BEEMD method achieves a closer match to the true velocity model across all depths from shallow to deep. This test using low-frequency-deficient data confirms that the proposed method effectively prevents the inversion from converging to local minima.

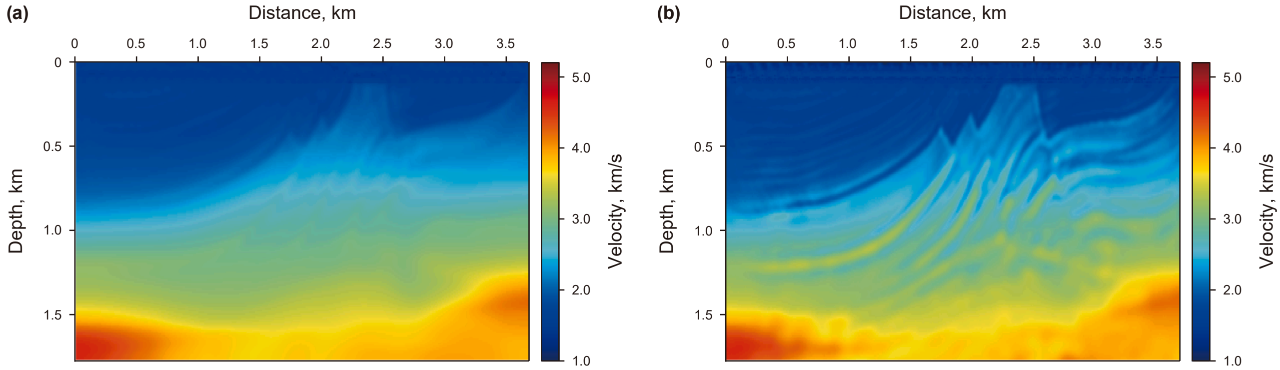

Fig. 27. FWI results after 100 iterations. The initial models of **(a)** and **(b)** are shown in Fig. 26(a) and (b), respectively.

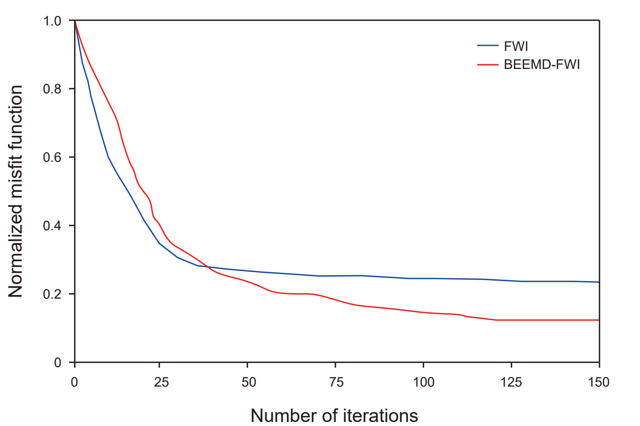

Fig. 28. Normalized misfit function convergence curves using different inversion methods.

4. Discussion

This study proposes a new FWI workflow that fully leverages both tomographic and migration gradients. Based on the tomographic and migration gradient, various inversion strategies have been proposed. The migration-mode gradient has been shown to enhance the accuracy of velocity model building and improve computational efficiency (Xie et al., 2024; Wang et al., 2024), while the tomography-mode gradient is effective in strengthening the robustness of the inversion process (Lian et al., 2018; Jeong et al., 2018; Hu et al., 2023).

The key innovation of this method lies in the BEEMD-based adaptive decomposition of the FWI gradient into tomographic and migration components, which enables a robust multi-scale inversion strategy. BEEMD achieves its efficacy through a noise-assisted ensemble averaging mechanism that effectively mitigates mode mixing and ensures reliable decomposition.

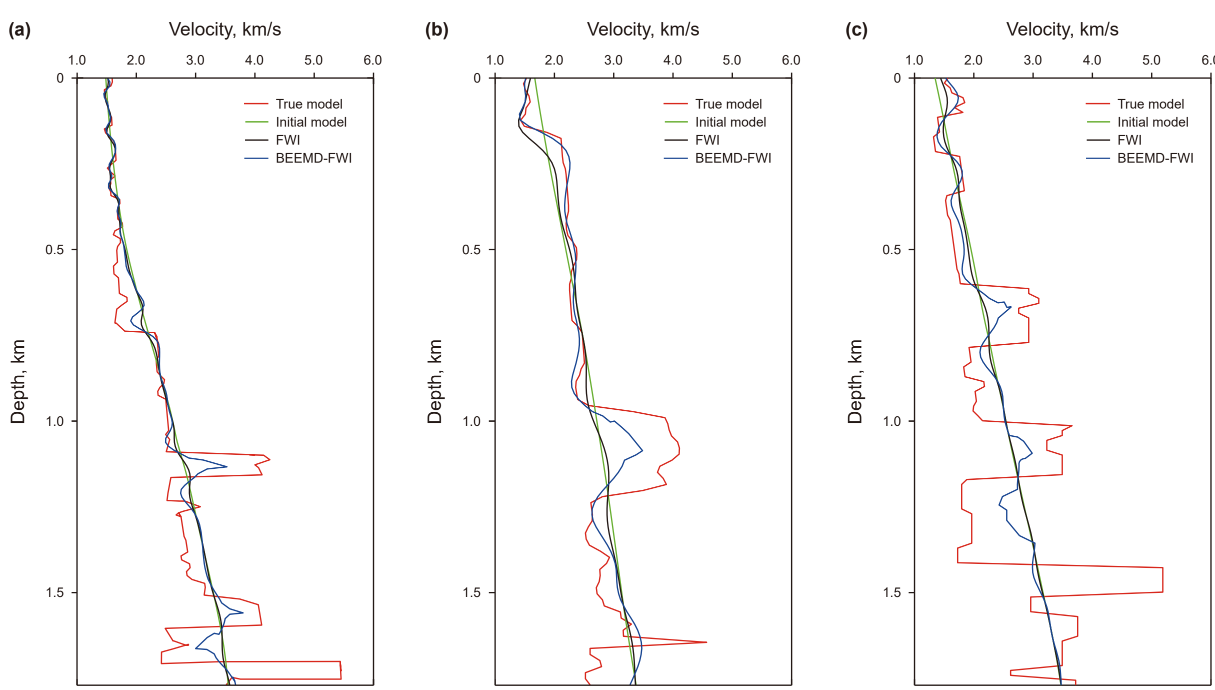

Fig. 29. Comparison of the velocity curves. These traces are extracted from the true velocity model, the initial velocity model, the conventional FWI result, and the BEEMD-FWI result at a distance of **(a)** 1.0 km, **(b)** 2.0 km, and **(c)** 3.0 km.

Table 2
Efficiency comparison between variable parameters using BEEMD-FWI.

Test number	σ(SD)	Number of decomposed BIMFs	N (ensemble number)	Computing time of one iteration, s	Percentage of computing time, %
1	0.1	2	50	3521	—
2	0.2	2	50	2403	68
3	0.1	4	50	6942	197
4	0.1	2	100	7212	204

However, the practical implementation of BEEMD requires careful consideration of two critical aspects: noise amplitude and computational efficiency. The noise amplitude (ε), which is the ratio of the standard deviations of the added noise to the original signal, governs the trade-off between scale separation and noise contamination. While insufficient noise (low ε) fails to prevent mode mixing, excessive noise necessitates a larger ensemble number (N) to average out artifacts and consequently raises the computational cost. Based on the analysis, we identified $\varepsilon = 0.2$ as an optimal balance. As shown in Table 2, computational efficiency is governed primarily by the stopping criterion (SD), the number of BIMFs, and N, with processing time scaling near-linearly with these parameters. To balance accuracy and practicality, we adopted a configuration of SD = 0.1, two BIMFs, and $N = 50$.

Computational efficiency and memory requirements are critical factors affecting the extension of this method to 3D. The interpolation of envelope surfaces in 3D space is inherently computationally expensive. Moreover, the unique ensemble averaging mechanism of BEEMD necessitates repeating this process dozens of times, which inevitably leads to a significant increase in computation time. The computation requires the simultaneous storage of the 3D gradient volume, intermediate iteration data, and multiple decomposed IMFs. This imposes stringent demands on the memory capacity of the computing nodes. Looking ahead, implementing 3D gradient decomposition must rely on efficient acceleration strategies to overcome these challenges. On the one hand, GPU-based parallel computing can accelerate interpolation and ensemble averaging while MPI domain decomposition divides large-scale models into sub-domains, thus reducing the memory load on individual nodes. On the other hand, a “2.5D” strategy can be adopted, where the 3D data volume is treated as a series of independent 2D slices along the survey lines. This approach significantly lowers the computational threshold while retaining the primary structural features, thus enhancing the feasibility of processing real-world 3D field data.

5. Conclusions

In this paper, we propose a strategy that uses the BEEMD method to decompose the conventional FWI gradient into migration and tomography components. BEEMD effectively decomposes the gradient into subcomponents of different scales, where the small-scale BIMFs are integrated as the migration component, and the large-scale residual component is regarded as the tomography component. When low-frequency data are abundant, the migration gradient dominates the inversion by refining high-wavenumber details. Conversely, when such data are lacking, the tomography gradient takes precedence to update the low-wavenumber background. Numerical experiments demonstrate that this method can effectively separate the components, not only improving the inversion accuracy but also providing a robust initial model for conventional FWI.

CRediT authorship contribution statement

Yi-Lin Lu: Writing – original draft, Methodology, Formal analysis. **Jian-Ping Huang:** Writing – review & editing. **Liang Chen:** Writing – review & editing. **Guo-Long Li:** Conceptualization. **Qiu-Yang Wang:** Writing – review & editing.

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

We would like to thank the editor and the reviewers for their valuable suggestions to this paper.

References

Alkhalifah, T., 2014. Scattering-angle based filtering of the waveform inversion gradients. Geophys. J. Int. 200 (1), 363–373. https://doi.org/10.1093/gji/ggu379.
Chen, L., Huang, J.P., 2024. Enhancing tomography component of full-waveform inversion based on gradient decomposition. IEEE Trans. Geosci. Rem. Sens. 62, 1–11. https://doi.org/10.1109/TGRS.2024.3456557.
Chi, B.X., Dong, L.G., Liu, Y.Z., 2014. Full waveform inversion method using envelope objective function without low frequency data. J. Appl. Geophys. 109, 36–46. https://doi.org/10.1016/j.jappgeo.2014.07.010.
Chi, B., Gao, K., Huang, L., 2017. Least-squares reverse time migration guided full-waveform inversion. In: 87th Annual International Meeting, SEG Expanded Abstracts, pp. 1471–1475. https://doi.org/10.1190/segam2017-17775494.1.
Fei, T.W., Luo, Y., Yang, J., et al., 2015. Removing false images in reverse time migration: the concept of de-primary. Geophysics 80, 237–244. https://doi.org/10.1190/geo2015-0289.1.
Fu, H., Yang, L., Miao, X., 2023. An efficient plug-and-play regularization method for full waveform inversion. J. Geophys. Eng. 20, 1140–1149. https://doi.org/10.1093/jge/gxad073.
Hu, G., Xu, W., He, B., et al., 2023. DIW-based reflection full waveform inversion and its application of land seismic data. J. Geophys. Eng. 20, 1109–1116. https://doi.org/10.1093/jge/gxad070.
Huang, N.E., Shen, Z., Long, S.R., et al., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. Proc. R. Soc. A A. 454 (1971), 903–995. https://doi.org/10.1098/rspa.1998.0193.
Huang, N.E., Shen, Z., Long, S.R., 1999. A new view of nonlinear water waves: the Hilbert spectrum. Annu. Rev. Fluid Mech. 31 (1), 417–457. https://doi.org/10.1146/annurev.fluid.31.1.417.
Huang, J.P., Qu, Y.M., Li, Q.Y., et al., 2015. Variable-coordinate forward modeling of irregular surface based on dual-variable grid. Appl. Geophys. 12, 101–110. https://doi.org/10.1007/s11770-014-0476-2.
Irabor, K., Warner, M., 2016. Reflection FWI. In: 86th Annual International Meeting, SEG Expanded Abstracts, pp. 1136–1140. https://doi.org/10.1190/segam2016-13944219.1.
Jeong, W., Tsingas, C., Kim, Y.S., 2018. Full waveform inversion with angle dependent gradient preconditioning using wavefield decomposition. J. Appl. Geophys. 159, 23–32. https://doi.org/10.1016/j.jappgeo.2018.07.016.
Kazei, V., Ovcharenko, O., Plotnitskii, P., et al., 2021. Mapping seismic data cubes to vertical velocity profiles by deep learning: new full-waveform inversion paradigm? Geophysics 86 (5), R1–R50. https://doi.org/10.1190/geo2019-0473.1.
Li, Y.Y., Alkhalifah, T., 2020. Multi-parameter reflection waveform inversion for acoustic transversely isotropic media with a vertical symmetry axis. Geophys. Prospect. 68 (6), 1878–1892. https://doi.org/10.1111/1365-2478.12966.
Lian, S.J., Yuan, S.Y., Wang, G.C., et al., 2018. Enhancing low-wavenumber components of full-waveform inversion using an improved wavefield decomposition method in the time-space domain. J. Appl. Geophys. 157, 10–22. https://doi.org/10.1016/j.jappgeo.2014.07.010.
Liu, F., Zhang, G., Morton, S.A., et al., 2011. An effective imaging condition for reverse-time migration using wavefield decomposition. Geophysics 76 (1), S29–S39. https://doi.org/10.1190/1.3533914.
Liu, D.J., Huang, J.P., Wang, Z.Y., 2020. Convolution-based multi-scale envelope inversion. Pet. Sci. 17 (2), 352–362. https://doi.org/10.1007/s12182-019-00419-8.
Métivier, L., Allain, A., Brossier, R., et al., 2018. Optimal transport for mitigating cycle skipping in full-waveform inversion: a graph-space transform approach. Geophysics 83 (5), R51–R68. https://doi.org/10.1190/geo2017-0807.1.
Mora, P., 1989. Inversion = migration + tomography. Geophysics 54 (12), 1575–1586. https://doi.org/10.1190/1.1442625.
Nie, Z., Feng, B., Wang, H., et al., 2025. Multi-scale full waveform inversion based on gradient bidimensional empirical mode decomposition. Geophysics 90 (5), 1–88. https://doi.org/10.1190/geo2024-0267.1.

Oristaglio, M., 2015. SEAM update: the Arid model—Seismic exploration in desert terrains. Lead. Edge 34 (4), 466–468. https://doi.org/10.1190/tle34040466.1.
Plessix, R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. Geophys. J. Int. 167 (2), 495–503. https://doi.org/10.1111/j.1365-246X.2006.02978.x.
Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain, Part 1: theory and verification in a physical scale model. Geophysics 64 (3), 888–901. https://doi.org/10.1190/1.1444597.
Ren, Z.M., Li, Z.C., Gu, B.L., 2019. Elastic reflection waveform inversion based on the decomposition of sensitivity kernels. Geophysics 84 (2), R235–R250. https://doi.org/10.1190/geo2018-0220.1.
Sun, D., Jiao, K., Cheng, X., et al., 2017. Born modeling based adjustive reflection full waveform inversion. In: 87th Annual International Meeting, SEG Expanded Abstracts, pp. 1460–1465. https://doi.org/10.1190/segam2017-17664306.1.
Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation. Geophysics 49 (8), 1259–1266. https://doi.org/10.1190/1.1441754.
Virieux, J., Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics. Geophysics 74 (6), WCC1–WCC26. https://doi.org/10.1190/1.3238367.
Wang, F., Donno, D., Chauris, H., et al., 2016. Waveform inversion based on wavefield decomposition. Geophysics 81 (6), R457–R470. https://doi.org/10.1190/geo2015-0340.1.
Wang, Z.Y., Huang, J.P., Liu, D.J., et al., 2019. 3D variable-grid full-waveform inversion on GPU. Pet. Sci. 16, 1001–1014. https://doi.org/10.1007/s12182-019-00368-2.
Wang, Z.Y., Wang, J., Sun, W., et al., 2024. Waveform inversion with structural regularizing constraint based on gradient decomposition. J. Geophys. Eng. 21 (4), 1119–1137. https://doi.org/10.1093/jge/gxae057.
Wu, Z., Alkhalifah, T., 2017. Efficient scattering-angle enrichment for a nonlinear inversion of the background and perturbations components of a velocity model. Geophys. J. Int. 210 (3), 1981–1992. https://doi.org/10.1093/gji/ggx283.
Wu, Z., Huang, N.E., 2009. Ensemble empirical mode decomposition: A noise-assisted data analysis method. Adv. Adapt. Data Anal. 1 (1), 1–41. https://doi.org/10.1142/S1793536909000047.
Xie, C., Qin, Z.L., Wang, J.H., Song, P., et al., 2024. Full waveform inversion based on hybrid gradient. Pet. Sci. 21 (3), 1660–1670. https://doi.org/10.1016/j.petsci.2024.01.013.
Xu, S., Wang, D., Chen, F., et al., 2012. Inversion on reflected seismic wave. In: 82nd Annual International Meeting, SEG Expanded Abstracts, pp. 1–7. https://doi.org/10.1190/segam2012-1473.1.
Xu, W.C., Hu, G.H., He, B.H., et al., 2022. Seismic reflection waveform inversion based on Gauss–Newton optimization. J. Geophys. Eng. 19 (4), 846–862. https://doi.org/10.1093/jge/gxac053.
Yang, J.D., Zhu, H.J., Li, X.Y., et al., 2020. Estimating P wave velocity and attenuation structures using full waveform inversion based on a time domain complex valued viscoacoustic wave equation: the method. J. Geophys. Res. Solid Earth 125 (6), e2019JB019129. https://doi.org/10.1029/2019JB019129.
Yao, G., Wu, D., 2017. Reflection full waveform inversion. Sci. China Earth Sci. 60, 1783–1794. https://doi.org/10.1007/s11430-016-9091-9.
Yao, G., da Silva, N.V., Warner, M., et al., 2018. Separation of migration and tomography modes of full-waveform inversion in the plane wave domain. J. Geophys. Res. Solid Earth 123 (2), 1486–1501. https://doi.org/10.1002/2017JB015207.
Yao, G., da Silva, N.V., Kazei, V., et al., 2019a. Extraction of the tomography mode with nonstationary smoothing for full-waveform inversion. Geophysics 84 (4), R527–R537. https://doi.org/10.1190/geo2018-0586.1.
Yao, G., da Silva, N.V., Warner, M., et al., 2019b. Tackling cycle skipping in full-waveform inversion with intermediate data. Geophysics 84 (3), R411–R427. https://doi.org/10.1190/geo2018-0096.1.
Yao, G., Wu, D., Wang, S.X., 2020. A review on reflection-waveform inversion. Pet. Sci. 17 (1), 334–351. https://doi.org/10.1007/s12182-020-00431-3.
Yong, P., Liao, W., Huang, J.P., 2018. Total variation regularization for seismic waveform inversion using an adaptive primal dual hybrid gradient method. Inverse Probl. 34 (4), 045006. https://doi.org/10.1088/1361-6420/aaaf8e.
Yong, P., Huang, J.P., Li, Z., et al., 2019. Least-squares reverse time migration via linearized waveform inversion using a Wasserstein metric. Geophysics 84 (8), S411–S423. https://doi.org/10.1190/geo2018-0619.