Petroleum Science 22 (2025) 1929–1945

Contents lists available at ScienceDirect

Petroleum Science

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

Original Paper

An EDCC-EMD analysis-based network for DAS VSP data denoising in frequency domain

Huan-Huan Tang [a, b], Shi-Jun Cheng [c], Wu-Qun Li [a, b], Wei-Jian Mao [a, b, *]

[a] Center for Computational & Exploration Geophysics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, 430077, Hubei, China
[b] State Key Laboratory of Precision Geodesy, Wuhan, 430077, Hubei, China
[c] Division of Physical Science and Engineering, King Abdullah University of Science and Technology, Thuwal, 23955-6900, Saudi Arabia

ARTICLE INFO

Article history:
Received 26 June 2024
Received in revised form
7 January 2025
Accepted 4 March 2025
Available online 6 March 2025

Edited by Meng-Jiao Zhou

Keywords:
DAS
Random noise
Coupling noise
EDCC-EMD
Deep learning

ABSTRACT

Distributed acoustic sensing (DAS) has rapidly emerged as a transformative technology in seismic exploration, particularly in vertical seismic profiles (VSP). However, the acquired VSP data suffer from strong coherent DAS coupling noise and random noise. Current deep learning denoising methods, dependent on noise labels derived from conventional denoising techniques, fall short in addressing the unique noise properties inherent in DAS data. To address this challenge, we propose an exponential decay curve-constrained empirical mode decomposition (EDCC-EMD) analysis-based supervised denoising network. Our method begins with extracting the initial noise from the field DAS VSP data through the traditional EMD method. Despite containing some signal leakage, this noise is further processed through EMD to derive intrinsic mode functions (IMFs). We, then, analyze the correlation coefficients between these IMFs and the initial noise, applying an exponential decay curve (EDC) law to isolate pure noise. This refined noise data serves as accurate labels, enhancing the denoising network's precision. Meanwhile, most of the methods usually consider the *t-x* domain features and ignore the important frequency-domain features. Consequently, we train our network with frequency-domain data instead of time domain data, capitalizing on the more distinct separation of noise and signal characteristics, thereby facilitating more effective noise-signal discrimination. The experimental results demonstrate that our method significantly enhances the denoising performance and successfully recovers weak signals.

1. Introduction

Distributed acoustic sensing (DAS) technology has been applied and developed rapidly in the field of vertical seismic profile (VSP) due to its advantages such as full borehole coverage, high-density layout, adaptability to high temperature and high-pressure environments, long service life, reusability, and low cost (Miller et al., 2012; Mateeva et al., 2014). However, due to the sensitivity and the poor coupling between the optical fiber cable and the casing, DAS-VSP data inevitably contain a significant amount of random noise and coherent DAS coupling noise (Yu et al., 2016; Chen et al., 2019). This substantial noise leads to severe signal contamination, and thus, results in low migration quality and reservoir prediction accuracy. Hence, removing random noise and DAS coupling noise is essential for improving the data's signal-to-noise ratio (SNR).

Seismic data denoising is a hot topic in the seismology community, and various methods have been developed for this purpose. These denoising methods can generally be categorized into random and coherent noise suppression based on the noise characteristics. For random noise, the techniques that utilize the difference of morphological characteristics of random noise and signals in the spatiotemporal domain (the former is a random distribution, while the latter is continuous), such as coherent filtering (Alsdorf, 1997; Li et al., 1997; Schimmel and Gallart, 2007), median filtering (Liu et al., 2006; Wang et al., 2020), structure-oriented space-varying median filter (Gan et al., 2016), and advanced median filter (Oboué et al., 2024), effectively remove random noise. Similarly, seismic data can be processed through the

* Corresponding author.
E-mail address: wjmao@whigg.ac.cn (W.-J. Mao).

https://doi.org/10.1016/j.petsci.2025.03.006

slant stack (Thorson and Claerbout, 1985) or sparse transformations, such as wavelet transform (Cai et al., 2001), optimal empirical wavelet transform (Geetha and Hota, 2023), curvelet transform (Zhang et al., 2018; Yin et al., 2024), and Radon transform (Kabir and Verschuur, 1995; Xue et al., 2017). In the transformed domain, the energy of the signals is strong while that of random noise is weak, allowing for differentiation by setting thresholds. Additionally, some studies have proposed incorporating matrix rank reduction theory into random noise suppression (Anvari et al., 2020; Oboué et al., 2024), based on which adaptive damped rank reduction method is proposed (Oboué et al., 2023). Moreover, in the frequency domain, random noise is distributed across the entire frequency band, while signals are confined to a limited bandwidth. Therefore, band-pass and FK filtering have been proven effective in suppressing random noise outside this finite bandwidth (Stein and Bartley, 1983; Binder et al., 2020; Xu et al., 2024). Overall, random noise is a type of noise that can be relatively easily removed, but there are still some challenges. For example, how to quickly select the optimal lengths of the time windows, thresholds, rank reduction parameters, and frequency band range for processing massive and complex seismic data. These parameter settings usually rely on manual experience and cannot be adjusted adaptively for complex data, often resulting in incomplete suppression of random noise or leakage of signals.

In the realm of coherent noise, various filtering methods can be employed to suppress it, based on the differences in the spectrum, apparent velocity, and arrival time between the coherent noise and the signals. For example, ground roll, which has lower frequencies and velocities compared to reflected signals, can be suppressed using a dip filter in the FK domain (Stein and Bartley, 1983; Anderson and McMechan, 1988; Taner et al., 1995; Binder et al., 2020). Also, this method is effective in suppressing refracted waves. To address the case that the coherent noise highly overlaps with the signal, local nonlinear filtering is proposed (Yuan et al., 2022). For multiples, according to the similarities of the multiple attenuation methods, these methods can be divided into two categories: filtering and predictive subtraction methods. The filtering methods mainly rely on the separability of primaries and multiples, including their periodicity and velocity differences. The most common methods in this category include predictive deconvolution (Taner et al., 1995) and Radon transform (Foster and Mosher, 1992; Wilson and Guitton, 2007; Song et al., 2022). The basis of the predictive subtraction methods is to use the wave equation to predict the multiples, and then, subtract them from the original data. In which, the most maturely developed and applied methods are the surface-related multiple elimination (SRME) (Dragoset et al., 2010; Kostov et al., 2015) and Marchenko equation-based methods (Alberto et al., 2017; Staring et al., 2020). However, the differences between coherent noise and signals are so subtle that it is hard to distinguish them in complicated seismic data. As a result, it may be difficult to suppress coherent noise effectively.

Recently, the rapidly developing field of deep learning (DL) technology has introduced new alternatives to seismic denoising (Zhang et al., 2020; LeCun et al., 2015; Jain and Seung, 2008; Birnie et al., 2021; Saad and Chen, 2020; Cheng et al., 2024). Compared to traditional model-driven denoising methods, DL methods have two distinct advantages. The first is their ability to learn automatically, independent of prior knowledge and manual experience (Zhang et al., 2017; Ross et al., 2018; Li et al., 2019). The second is that, based on big data, DL methods can extract multi-dimensional and nonlinear deep features of seismic data (Zhang et al., 2019; Yang et al., 2020). Leveraging these two advantages, DL algorithms can overcome the limitations of traditional denoising methods, which overly rely on manual experience, and are anticipated to exhibit substantial potential in seismic denoising (Zhang et al., 2020; Birnie et al., 2021; Jin et al., 2018). DL-based denoising methods can be classified into supervised learning (SL) (Mandelli et al., 2019; Yu et al., 2019; Dong et al., 2020; Feng and Li, 2021), self-supervised learning (SSL) (Shao et al., 2022; Wang et al., 2023c; Cheng et al., 2024a,b; Liu et al., 2024), and unsupervised learning (Song et al., 2024; Wang et al., 2023a; Zhang and Wang, 2023; Yang et al., 2023; Qian et al., 2024), depending on whether model training requires labeled data. The denoising capability of SL methods is closely related to the labeled data, which is often challenging to obtain in real-world scenarios. The SSL methods rely on the pseudo-labels obtained through the training data itself, while unsupervised methods typically utilize generative models to learn the characteristics of signals without the labeled data. Currently, the last two kinds of methods are mainly used to suppress random noise which is easier to remove (Wang et al., 2023a; Zhang and Wang, 2023).

For DAS VSP data, which is the focus here, both random noise and coherent coupling noise coexist. The energy of random noise is extremely strong, and the coupling noise blends with the signals at certain frequencies (Zhang et al., 2020). This complexity makes it challenging for traditional denoising methods to remove all noise without signal leakage. Meanwhile, unsupervised and SSL approaches also struggle to identify both types of noises concurrently. This is because they learn signal characteristics without labeled data, and therefore, might not effectively distinguish between random noise and coherent coupling noise in complex data that contain both. In this context, SL stands out as a superior denoising method for DAS VSP data, and building on SL, scholars have pursued research in two directions. First, trained in the *t-x* domain, researchers constructed different neural networks to suppress seismic noise, such as, Chen et al. (2018) proposed GAN-CNN based Blind Denoiser (GCBD), which employs adversarial learning to extract noise for supervised learning. One limitation of GCBD is that the noise is assumed to be additive noise with zero mean (Chen et al., 2018), while the noise in DAS VSP data consists of random noise that follows a Gaussian distribution with zero mean, as well as coupling noise with more complex characteristics that do not meet these criteria (Feng et al., 2024). Jin et al. (2018) proposed an auto encoder built upon the deep residual network to perform noise attenuation; Zhao et al. (2018) modified the original DnCNN from the aspects of patch size, convolution kernel size, network depth, and training set to suppress the non-Gaussian random noise; Richardson and Feller (2019) used deep learning with a U-Net model incorporating a ResNet architecture to denoise seismic data; Saad and Chen (2020) proposed to attenuate random noise based on a deep-denoising autoencoder. Second, trained in the transformed domain; Feng and Li (2021) trained a network with singular spectrum instead of the time domain data to suppress DAS noise, which can represent geophysical features more accurately and help separate signals from noises; Iqbal (2022) proposed DeepSeg working on segments of the time-frequency domain, which automatically determines the signal information contained by each data point in the time-frequency plane. In this method, the seismic trace is first transformed into the 2-D domain, and the network's input is still a single-channel signal, for two-dimensional DAS VSP data, it still has to be processed trace by trace, which is not applicable to the DAS VSP record with a huge data volume. Wang et al. (2023b) proposed using multiple 2D discrete wavelet transform (2D-DWT) to convert labels in the *t-x* domain into the DWT domain. The DWT can filter and down-sample the data along the horizontal and vertical directions, which can effectively improve the training and prediction efficiency. For seismic data denoising, the establishment of a general intelligent denoising network is a shared objective. However, before achieving this goal, it is both meaningful and necessary to conduct comprehensive and

meticulous research including various neural networks as well as training datasets with different domains.

In our manuscript, we analyze the characteristics of signals, random noise, and coupling noise in the frequency domain. We discover that the signals are distributed within a limited bandwidth and are continuous in the spatial direction; the random noise is distributed across the entire frequency band with weaker amplitude; the coupling noise is distributed in a few narrow frequency bands and is discontinuous in the spatial direction. These characteristics are very conducive for the network to distinguish them. Therefore, we train the network with *f-x* domain data other than the *t-x* domain data, aiding the network in more easily distinguishing between noise and signals. Additionally, we proposed an EDCC-EMD noise extraction method based on the traditional EMD. Based on EDCC-EMD, we can swiftly and accurately extract pure random noise and coupling noise from field DAS VSP data. The extracted noise is injected into synthetic data to construct a corresponding training dataset. This strategy helps reduce the feature gap between synthetic and field data, thereby enhancing the generalizability of our method on field data. Based on the excellent performance of U-Net in image denoising (Falk et al., 2019), U-Net is employed as our baseline network for denoising DAS VSP data. Finally, we test the performance of our method on both synthetic data and field data.

The main contributions of this study are summarized as follows: (1) We proposed the EDCC-EMD method for noise extracting, which effectively overcomes the difficulty in obtaining accurate labels for training in real-world scenarios. (2) The EDCC-EMD can be applied not only to DAS VSP data but also more broadly across fields, such as seismic data collected by conventional geophones or earthquake data. (3) The shift to frequency-domain training data enhances the network's learning efficiency and effectiveness in noise-signal separation. (4) The feasibility of the proposed method in denoising the random noise and coupling noise simultaneously is demonstrated by the field DAS VSP results.

2. Theory

In this section, we first review the traditional EMD denoising methods and their limitations in DAS VSP data denoising. Then, the EDCC-EMD method is introduced to extract the pure noise from field DAS VSP data. Next, the labeled data generation is shown. Finally, we present the denoising network architecture and the training configurations.

2.1. The traditional EMD denoising method

EMD, proposed by Huang et al. (1998), is a breakthrough technique for analyzing non-stationary signals and has been widely used for noise suppression of seismic data (Battista et al., 2007; Gómez and Velis, 2016; Bekara and van der Baan, 2009). It essentially stabilizes the unsteady signals by decomposing them into a series of intrinsic modal functions (IMFs) and a residual. Each IMF has a relatively constant local frequency, decreasing with the order of the IMF, and needs to satisfy two conditions: (1) In the whole data, the number of zero values and extreme values are equal or differ by at most one; (2) At any point, the mean value of the envelope defined by the local maxima and local minima is zero (Huang et al., 1998). This decomposition process is reversible, as the original signal can be obtained without information loss by summing all the IMFs and the residual. The decomposition can be expressed as follows:

$$x(t) = \sum_{i=1}^{M} c_i(t) + r(t), \tag{1}$$

where $x(t)$ is the original seismic data, $c_i(t)$ is the IMF decomposed, M is the number of IMFs, and $r(t)$ is the residual. EMD can adaptively decompose the nonlinear and unsteady signal features of seismic data into different frequency bands.

We, then, can calculate the correlation coefficient between each IMF and the original data as follows:

$$R(i) = \sum_{j=1}^{N} c_i(t_j) \cdot x(t_j) \Big/ \sum_{j=1}^{N} x(t_j) \cdot x(t_j). \tag{2}$$

In Eq. (2), $R(i)$ is the correlation coefficient between the IMF_i and the original data x, and N is the sample number along the time direction. With the maximum correlation coefficient $R(k)$, we can determine that the IMFs of order 1 to $k-1$ are noise and k to M are signals (Liu and Chen, 2016; Chen et al., 2017). The separated signals $s(t)$ and noise $n(t)$ can be represented as

$$s(t) = \sum_{i=k}^{M} c_i(t), \tag{3}$$

$$n(t) = \sum_{i=1}^{k-1} c_i(t). \tag{4}$$

Based on Eqs. (1)–(4), noise suppression or extraction in seismic data can be realized. However, due to the mode-mixing issue, the signals and noise will be blended during the decomposition process, which causes partial signal loss in the removed data and residual noise in the denoising result (Gómez and Velis, 2016; Chen et al., 2014, 2017). Fig. 1 shows the ability of EMD in seismic data denoising. The noisy data is shown in Fig. 1(a). We can observe that EMD can effectively remove noise and enhance the continuity of the events (Fig. 1(b)). However, it is worth noting that substantial signal leakage can be found in the removed noise profile (Fig. 1(c)). Hence, noise directly extracted from field data by EMD methods is not suitable as labels for DL methods. Instead, in the following, we will introduce an EDCC-EMD method to extract noise.

2.2. The EDCC-EMD noise extraction method

Based on the traditional EMD method, we can obtain the initial noise data with some signal leakage from the field DAS VSP data. Due to the mode mixing during the decomposition process, it is difficult to separate noise and signals with the same wave patterns using the EMD method. Hence, the initial noise data could be divided into two types: pure noise and noise mixed with leaked signals. The first objective of our work is on how to further select the pure noise from the initial noise data. The selection criterion of pure noise is to judge whether the correlation coefficient curve between the IMFs of noise and noise data satisfies the exponential decay curve (EDC), expressed as follows:

$$\begin{cases} \max(R_i) = R_1, \\ \sum_{i=2}^{M} R_i < \lambda R_1. \end{cases} \tag{5}$$

In Eq. (5), λ is normally chosen as $0 \le \lambda \le 0.1$. If the correlation coefficient curve $R(i)$ satisfies Eq. (5), then this trace is pure noise, otherwise, it is noise containing the leakage signal. This method is

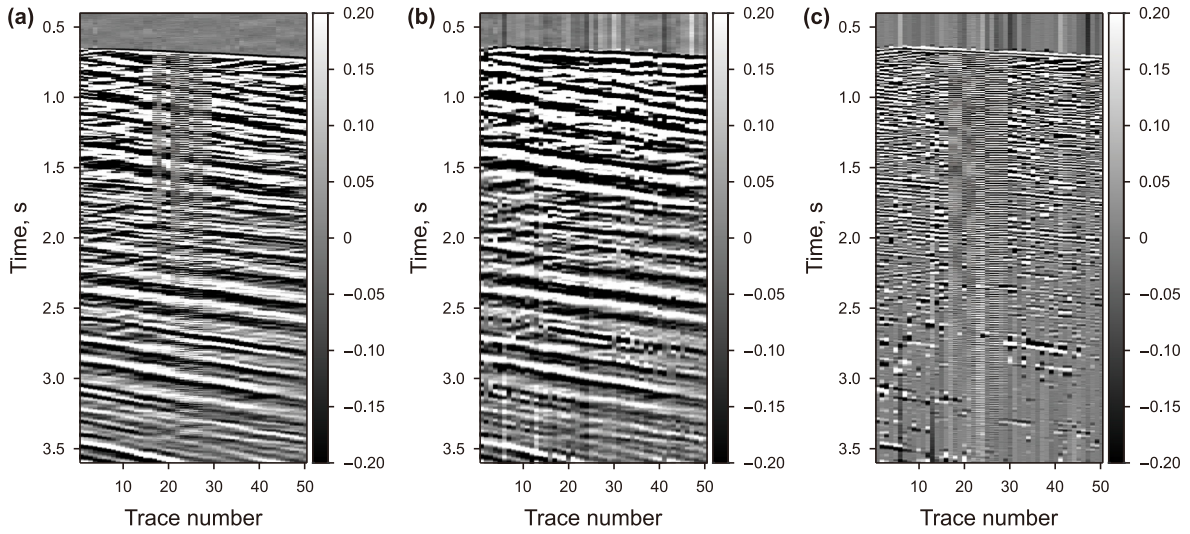

Fig. 1. Seismic data denoising by EMD. **(a)** Original data with coupling noise. **(b)** Denoised data. **(c)** Removed noise.

termed exponential decay curve-constrained empirical mode decomposition (EDCC-EMD). The detailed workflow of the EDCC-EMD is illustrated in Fig. 2.

The field DAS VSP data from a certain seismic survey area is shown in Fig. 3. The original acquired data from one optical fiber includes 214 shots spaced approximately 37.7 m. Each shot uses 324 receivers with a spacing of 10 m and has 2000 time samples with a time interval of 2 ms. We, here, show only 4 shot gathers of them. From different shot gathers in Fig. 3, we can see evident coupling noise, as indicated by the red arrows, which are distributed at the same receivers; as well as random noise, highlighted by the red rectangle. It should be noted that, unlike conventional seismic recording, due to the continuous monitoring of DAS, some weak signals from passive sources (such as micro-earthquakes) or the previous shot will be recorded before the first arrival of the current shot in the DAS VSP recording. Therefore, the data in the red rectangle contains both random noise and weak signals. Fig. 4(a) is the zoom view of the data in the red rectangle. We use both the EMD and EDCC-EMD to extract random noise and coupling noise from these 214 shot gathers. The initial random noise extracted via the EMD is shown in Fig. 4(b), while the pure random noise extracted from the initial random noise by the EDCC-EMD is presented in Fig. 4(c). It is evident that the initial random noise extracted by the EMD contains a small amount of leakage signal, as shown by the red arrow in Fig. 4(b), whereas the pure random noise extracted by the EDCC-EMD does not have any signal leakage.

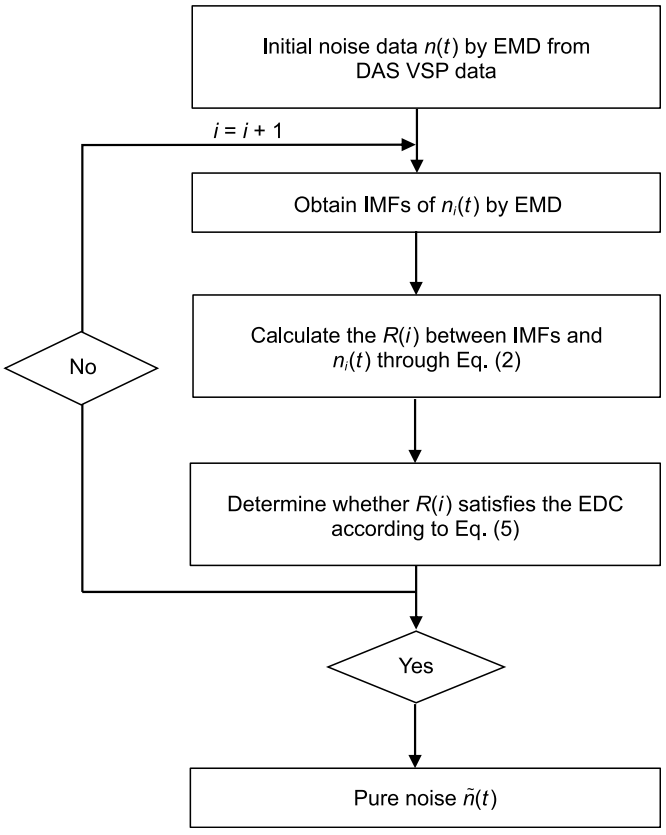

Fig. 2. The workflow of the EDCC-EMD used in noise extraction.

After extracting the noise, we further analyze the decomposed IMFs characteristics and correlation coefficient curves of the original data, initial random noise, and pure random noise, which is shown in Fig. 5. The noisy data of the 214th trace in Fig. 4(a) is shown in Fig. 5(a1), and the seven decomposed IMFs are respectively displayed in Fig. 5(a2)–(a8). The correlation coefficients between each IMF and the data in Fig. 5(a1) are illustrated in Fig. 5(a9). Similarly, the IMFs and their correlation coefficient curves $R(i)$ for initial random noise and pure random noise are shown in Fig. 5(b1)–(b9) and Fig. 5(c1)–(c9), respectively. Comparing the correlation coefficient curves between IMFs and the data itself, it can be seen that the maximum values of $R(i)$ for both types of extracted random noises gradually approach the low-order IMFs. For the initial random noise, the second-order IMF of the 214th trace still shows high similarity to the data, while for the pure random noise, only the first-order IMF is highly correlated with the data. The correlation coefficients of the pure random noise follow an exponential decay curve, which is consistent with the EDCC-EMD theory.

For coupling noise, in addition to Eq. (5) serving as the extraction principle, the consistent distribution position across different shot gathers also acts as a constraint condition. Fig. 6 displays a portion of the extracted initial coupling noise by EMD. As seen, we can identify the zig-zag coupling noise with strong energy, the chessboard-shaped coupling noise with weaker energy, and the

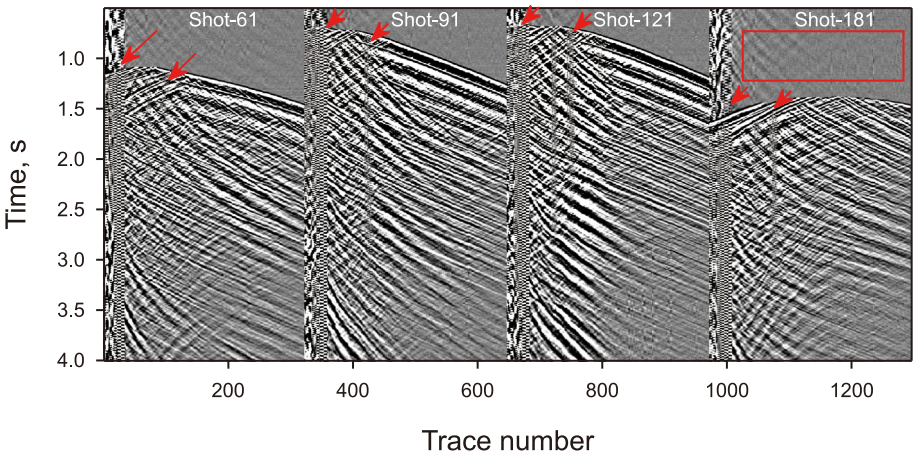

Fig. 3. The 4 shot gathers of field DAS VSP.

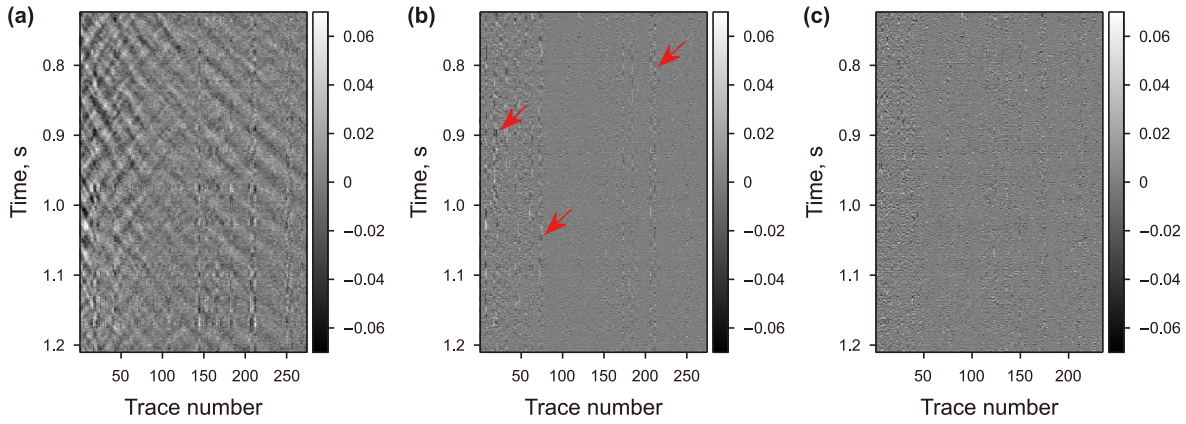

Fig. 4. The extracted random noise. **(a)** The original data in the red rectangle in Fig. 3. **(b)** Initial random noise extracted by EMD. **(c)** Pure random noise extracted by EDCC-EMD.

partially leaked continuous signal (shown by the red arrows).

The coupling noise trace-154 mixed with signals, its IMFs, and correlation coefficient curves $R(i)$ are shown in Fig. 7(a1)–(a7), respectively. We can see that $R(i)$ for trace-154 does not follow an exponential attenuation curve. The coupling noise trace-144, its IMFs, and correlation coefficient curves $R(i)$ are shown in Fig. 7(b1)–(b7), respectively. Obviously, the $R(i)$ of trace-144 is an exponential decay curve. The pure coupling noise we further extracted such as trace-144 by EDCC-EMD from the initial coupling noise, which is based on the characteristics of these correlation coefficient curves, is shown in Fig. 8(a). According to their distribution positions, they can be divided into zig-zag coupling noise (Fig. 8(b)) and checkerboard-shaped coupling noise (Fig. 8(c)). The time-frequency spectra of these two forms of coupling noise are shown in Fig. 8(d) and (e), respectively. From these figures, we can clearly see that the coupling noise is distributed across several narrow frequency bands, with its amplitude discontinuously distributed in energy clusters over time direction. The characteristics of the coupling noise in the frequency domain are quite different from those of the signals, which helps the DL models to distinguish them.

2.3. The training data generation

We use the simulated data, which is generated with a 2D profile derived from the SEAM, to constitute our training data set. Fig. 9 shows the velocity model used here, and the model has 70 km width with 10 m interval and 5 km depth with 5 m interval. We utilize the finite-difference algorithm to perform forward modeling. A Ricker wavelet with a dominant frequency ranged in 10–25 Hz is used the source signal. We generate a total of 176 shot gathers, each with the size of 3001×401, where 3001 is the number of time samples with 2 ms interval, and 401 is the number of receivers with 10 m interval. The first 100 shot gathers are used to make training data and the last 76 shot gathers are for testing.

Since the optimization of the neural network is conducted in an SL manner, our training dataset comprises both noisy inputs and corresponding labeled data. The label data is directly extracted from the simulated clean data. To reduce the feature gap between the synthetic and field data, the corresponding noisy input is obtained through the simulated data and the extract real noise by the proposed EDCC-EMD from the field data, shown as Eq. (6):

$$d_2 = d_1 + \alpha\hat{n}_1 + \beta\hat{n}_2 + \gamma\hat{n}_3, \tag{6}$$

where d_2 is the noisy data, d_1 is the simulated data, $\hat{n}_1$ is the extracted pure random noise, $\hat{n}_2$ is the pure zig-zag coupling noise, and $\hat{n}_3$ is the pure checkboard-shaped coupling noise. The parameters α, β, and γ are weights of noise respectively, which are used to adjust the level of added noise, thereby expanding the training data set to enhance the generalization ability of the network. Their values are randomly selected between $[0.001, 1]$. The clean data d_1 with the size 120×120 is patched from the shot gathers, and the amplitude of the shot gathers is normalized to

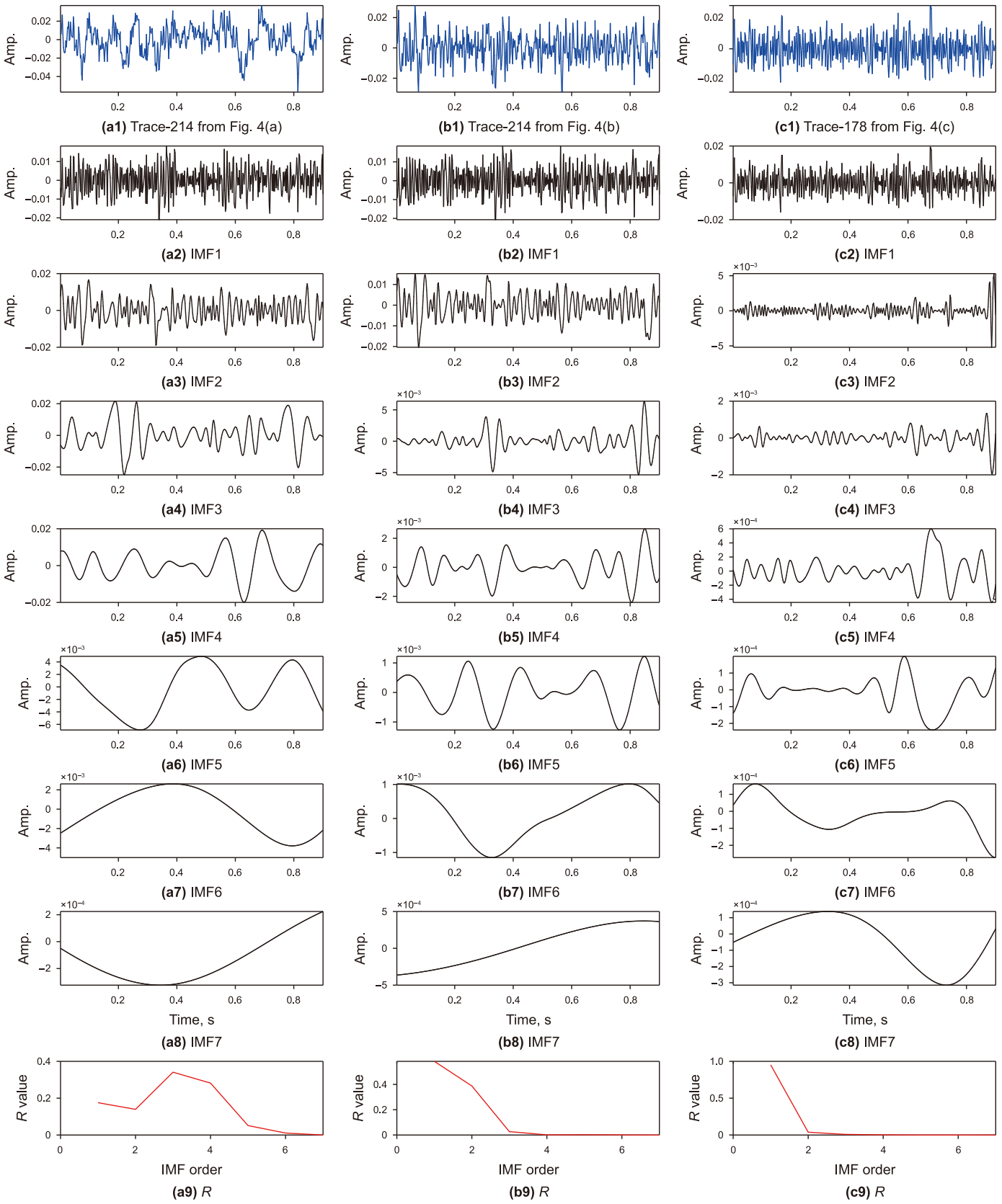

Fig. 5. Comparisons of the IMFs and the correlation coefficient curves of random noise. **(a1)** The 214th trace of the original data. **(a2)**–**(a8)** The 7 IMFs of the original data, respectively. **(a9)** The $R(i)$ of the original data. **(b1)** The 214th trace of the initial random noise by EMD. **(b2)**–**(b8)** The 7 IMFs of the initial random noise, respectively. **(b9)** The $R(i)$ of the initial random noise. **(c1)** The 178th trace of the pure random noise by EDCC-EMD. **(c2)**–**(c8)** The 7 IMFs of the pure random noise, respectively. **(c9)** The $R(i)$ of the pure random noise. The abbreviation Amp. indicates amplitude.

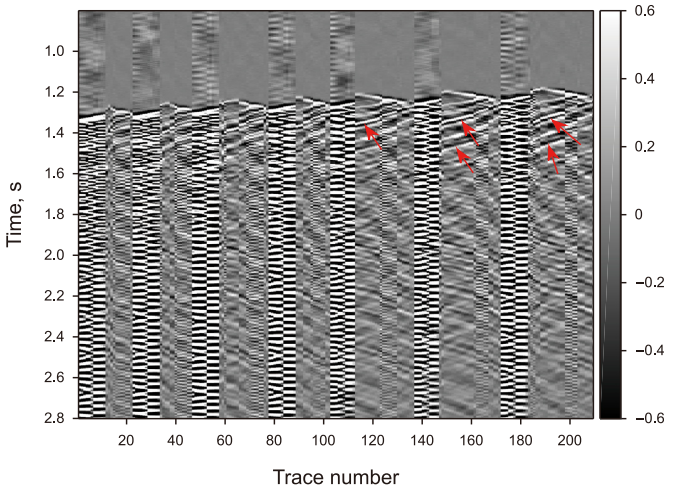

Fig. 6. Part of the extracted initial coupling noise by EMD.

$[-1, 1]$. The three different types of noises with the size 120×120 are all from the field data extracted via EDCC-EMD. Unlike the random noise $\widehat{n}_1$, only some traces are valid for the pure zig-zag coupling noise $\widehat{n}_2$ and the pure checkboard-shaped coupling noise $\widehat{n}_3$ in the 120 traces, the number of valid traces is randomly set from [0,40], and their positions are randomly distributed, and the other traces are invalid filled with 0 values. Here, we get a total of 40,428 pairs of noisy and clean data in the time domain.

Data augmentation techniques are employed such as patches, real noises, and varying SNRs to increase the diversity of the training data. This helps networkd better generalize to different noise conditions and signal types, enhancing the denoising effect. Moreover, based on the earlier findings that the features of signals and noise exhibit greater differences in the frequency domain, thus making them easier to distinguish, we opt for using frequency domain data instead of time domain data for training the network. This stagey is more conducive to network optimization, as we demonstrate later. The formula for transforming seismic data from the time domain to the frequency domain is as follows:

$$D(\omega)=\int_{-\infty}^{+\infty} d(t)\exp(-i\omega t)\mathrm{d}t. \tag{7}$$

In Eq. (7), $D(\omega)$ is the frequency domain data, $d(t)$ is the time domain data, ω is the frequency, and t is the time. Since here the real number domain U-Net is utilized, the complex domain data after Fourier transformation needs to be divided into real and imaginary parts for independent processing. Consequently, the 40428 pairs of training data in time domain are doubled to 80856 pairs in frequency domain, and some of them are shown in Fig. 10.

The random noise and coupling noise not only elevate the complexity of noise suppression but also set more stringent standards for the design and assessment of denoising algorithms. From the frequency spectra of signals, random noise, and coupling noise, the characteristics of their distribution can be found, as shown in Fig. 10. The signals are distributed within a limited bandwidth and are continuous in the spatial direction. The random noise is distributed across the entire frequency band with weaker amplitude. The coupling noise is distributed in a few narrow frequency bands and is discontinuous in the spatial direction. These characteristics are very conducive for the network to distinguish them.

2.4. The denoising network architecture

Based on the excellent performance of U-Net in image denoising (Zhang et al., 2019), we adopt U-Net as our baseline network to denoise DAS VSP data. The network architecture for DAS VSP denoising is shown in Fig. 11. The symmetric structure of U-Net, including the contracting and expansive paths, helps capture the contextual information of the signal. In denoising applications, this structure is beneficial to preserve useful information from the seismic signal while removing noise. Furthermore, U-Net's skip connections combine low-level detail information with high-level contextual information, which is crucial for differentiating noise from signal. In the frequency domain, this combination helps to identify and remove noise while preserving the spectral characteristics of the seismic signal. The input of the U-Net model is the frequency domain noisy data, and the label is the clean data in the frequency domain. In the frequency domain, the signals are distributed within a limited bandwidth and are continuous in the spatial direction. The random noise is distributed across the entire frequency band with weaker amplitude. The coupling noise is distributed in a few narrow frequency bands and is discontinuous in the spatial direction. Signals and noises exhibit different distribution characteristics in the frequency domain, which U-Net can utilize for effective feature discrimination. For the convenience of narration, we will use the abbreviation: "U-Net-*f*" to represent the network trained with the frequency domain data, and "U-Net-*t*" to represent that trained with the time domain data in the following. By optimizing the network parameters θ, the network f can learn the nonlinear mapping between the noisy data D_2 and the clean data D_1. We employ the mean square error as our loss function, which is defined as

$$L(\theta)=\frac{1}{Q}\sum_{j=1}^{Q}\left(f\left(\theta, {D_2}^j\right)-{D_1}^j\right)^2, \tag{8}$$

where Q is the number of samples in the training. The detailed training configuration settings are shown in Table 1. These configurations are important for the optimization of network, but are not the focus of this paper, so we refer to the relevant research (Zhang et al., 2017, 2019; Dong et al., 2020; Wang et al., 2023a).

We utilize the SNR to quantitatively measure the denoising ability, which has the form:

$$SNR=10\lg\frac{||d||_2^2}{||d-\widehat{d}||_2^2}, \tag{9}$$

where d represents the true clean data and $\widehat{d}$ represents the denoised data. For the computational cost of training the DAS VSP data denoising network, it takes 62 h 21 min on a GPU with an NVIDIA RTX 2080Ti in the environment of Pytorch (version 1.10.2) framework.

3. Experiments

In the following, we first share the denoising products of the proposed U-Net-*f* on both the synthetic and field DAS VSP data, and also, compare the results with those of three other denoising methods, including EMD method, FK method, and U-Net-*t*.

3.1. Synthetic data

For synthetic data, the last 76 shot gathers generated on SEAM model are used to demonstrate the effectiveness of the proposed

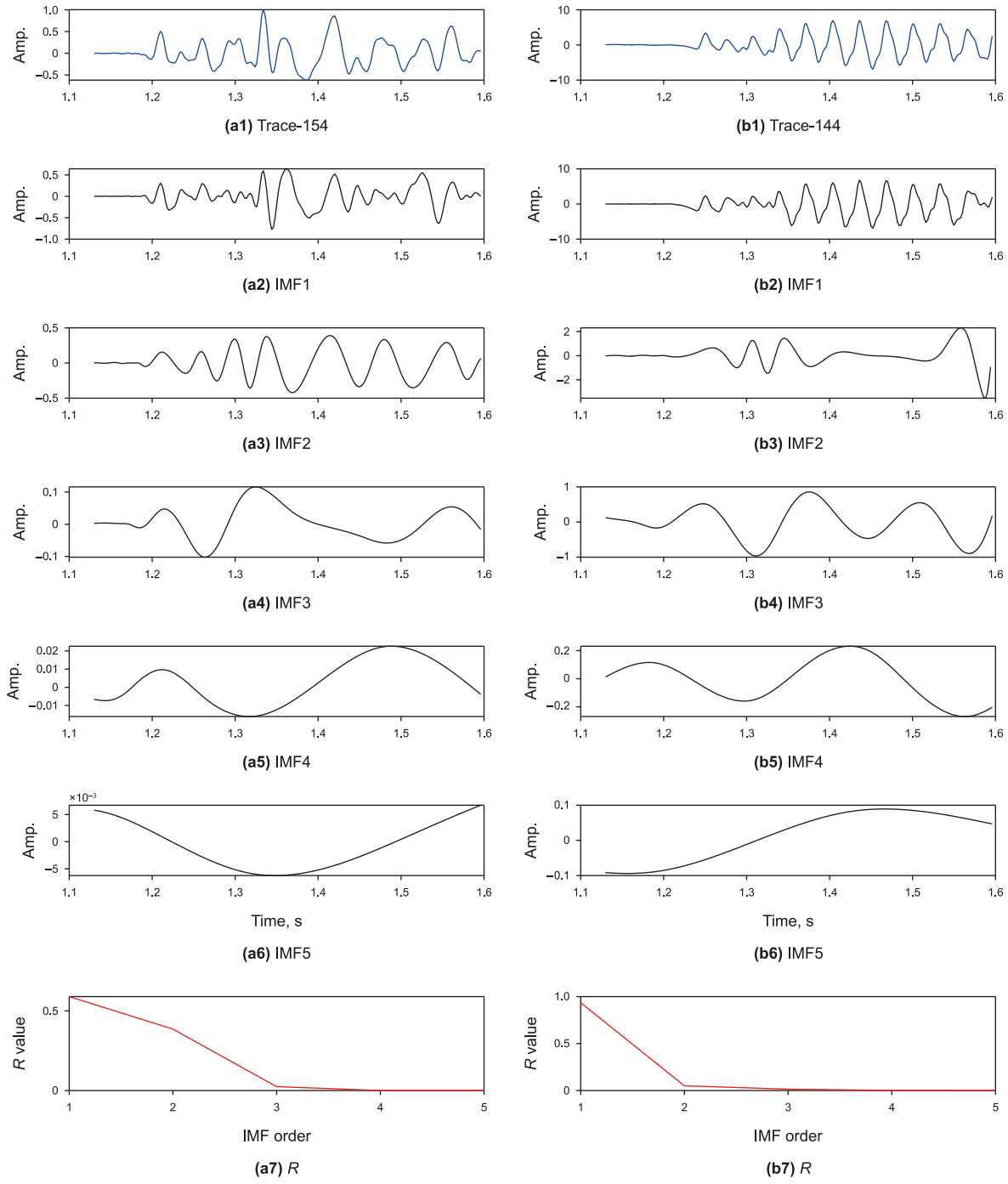

Fig. 7. Comparisons of the IMFs and the correlation coefficient curves of coupling noise. **(a1)** The 154th trace of the initial coupling noise. **(a2)**–**(a6)** The 5 IMFs of the 154th trace, respectively. **(a7)** The $R(i)$ of the 154th trace. **(b1)** The trace-144 of the initial coupling noise. **(b2)**–**(b6)** The 5 IMFs of the 144th trace, respectively. **(b7)** The $R(i)$ of the 144th trace.

method. To validate the denoising performance of our method, we add random noise and coupling noise to the clean data. Here, we show one of the test shot gathers (shot-125). The original clean data as the reference for denoising results is shown in Fig. 12(a), and the noisy data with 7.6 dB is shown in Fig. 12(b). Fig. 12 (c) and 12(d) are their FK spectrum respectively. The weak signal below 3 s is almost buried in the noise (see Fig. 12(b)), and it is hard to distinguish the signal.

Comparing the FK spectra of these two data, we can see that the strong coupling noise pollutes the signal. The frequencies around 30 Hz and 36 Hz, the coupling noise and the signal are blended, which increases the difficulty of traditional methods to separate them. The random noise is distributed in the full frequency band and has weaker amplitude on the FK spectra. The denoising results, obtained by EMD, FK, U-Net-*t*, and the U-Net-*f*, are shown in Fig. 13. In the following, we will compare and analyze the denoising products of the four methods from five aspects.

First, the denoised data of these four methods are shown in Fig. 13(a)–(d), respectively. We can see that, EMD, FK, and U-Net-*t* partially remove the random noise and coupling noise, with EMD leaving the most residual coupling noise. The FK method introduces some strong truncation noise near the first arrival. U-Net-*t* retains a

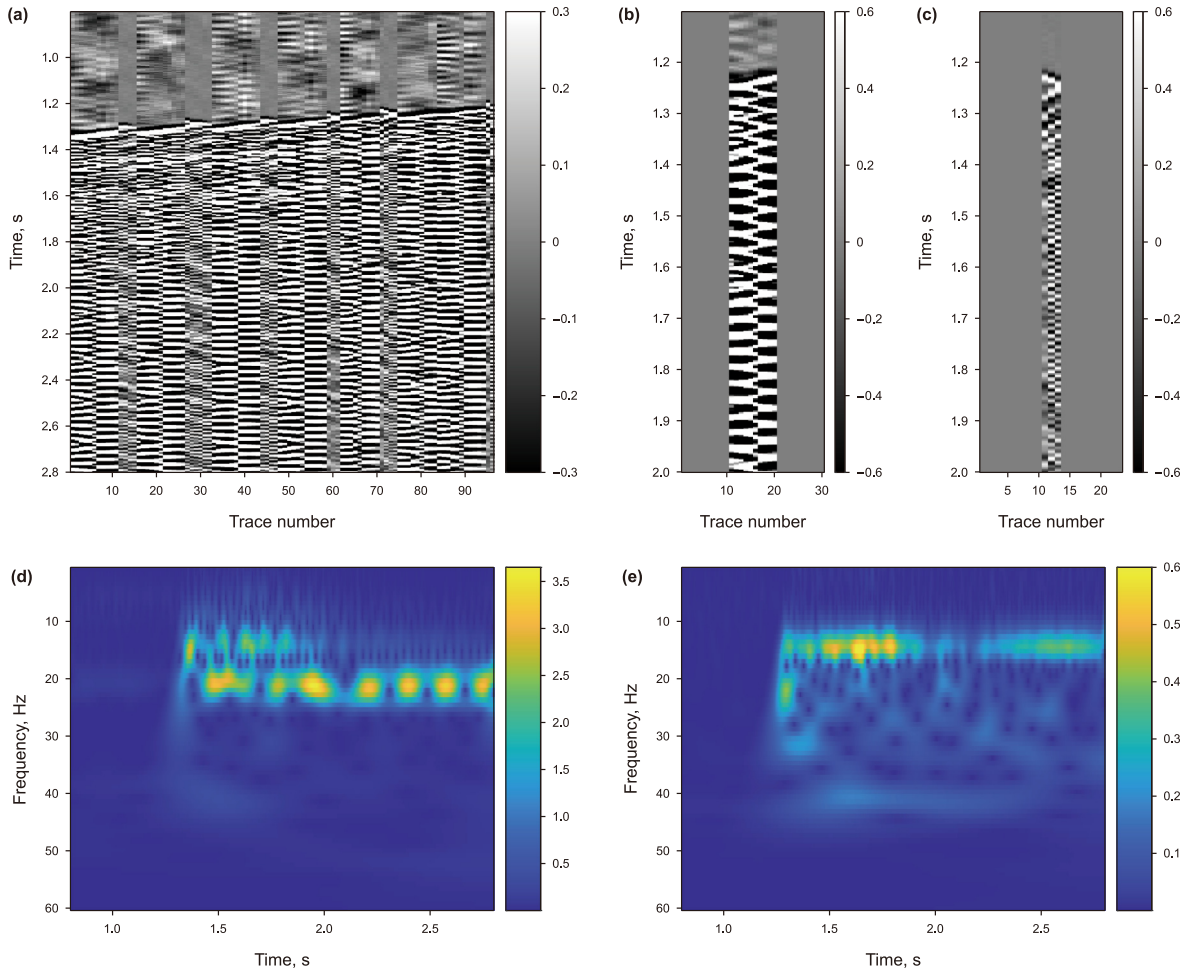

Fig. 8. The pure coupling noise by EDCC-EMD from the initial coupling noise and their time-frequency spectra. **(a)** The pure coupling noise from Fig. 6. **(b)** Zig-zag coupling noise. **(c)** Checkerboard-shaped coupling noise. **(d)** The time-frequency spectrum of the 11th trace in (b). **(e)** The time-frequency spectrum of the 11th trace in (c).

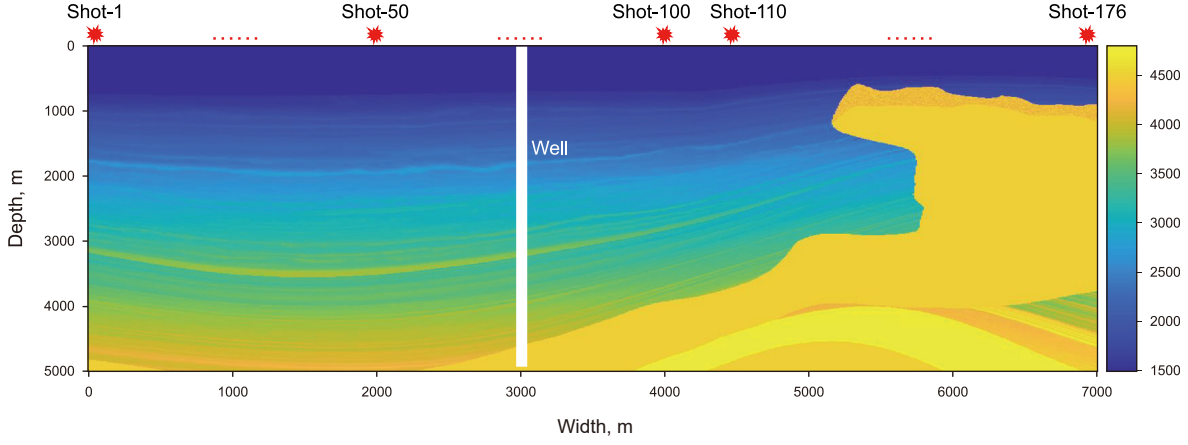

Fig. 9. Two-dimensional profile of the SEAM velocity model and the acquisition geometry.

small amount of coupling noise. In contrast, U-Net-*f* effectively eliminates both random noise and coupling noise and recovers the weak signals below 3 s.

Second, the removed noise by the four methods is presented in Fig. 13(e)–(h), respectively. It is evident that both EMD and U-Net-*t* methods tend to damage the signal integrity, while U-Net-*f* preserves signal amplitude well.

Third, the FK spectrum of the denoising results, presented in Fig. 13(i)–(l), respectively, shows the EMD method does not remove the low-frequency random noise. Also, we can see that EMD, FK, and U-Net-*t* methods struggle to suppress the coupling noise sharing the frequency with signals. Conversely, U-Net-*f*, through its

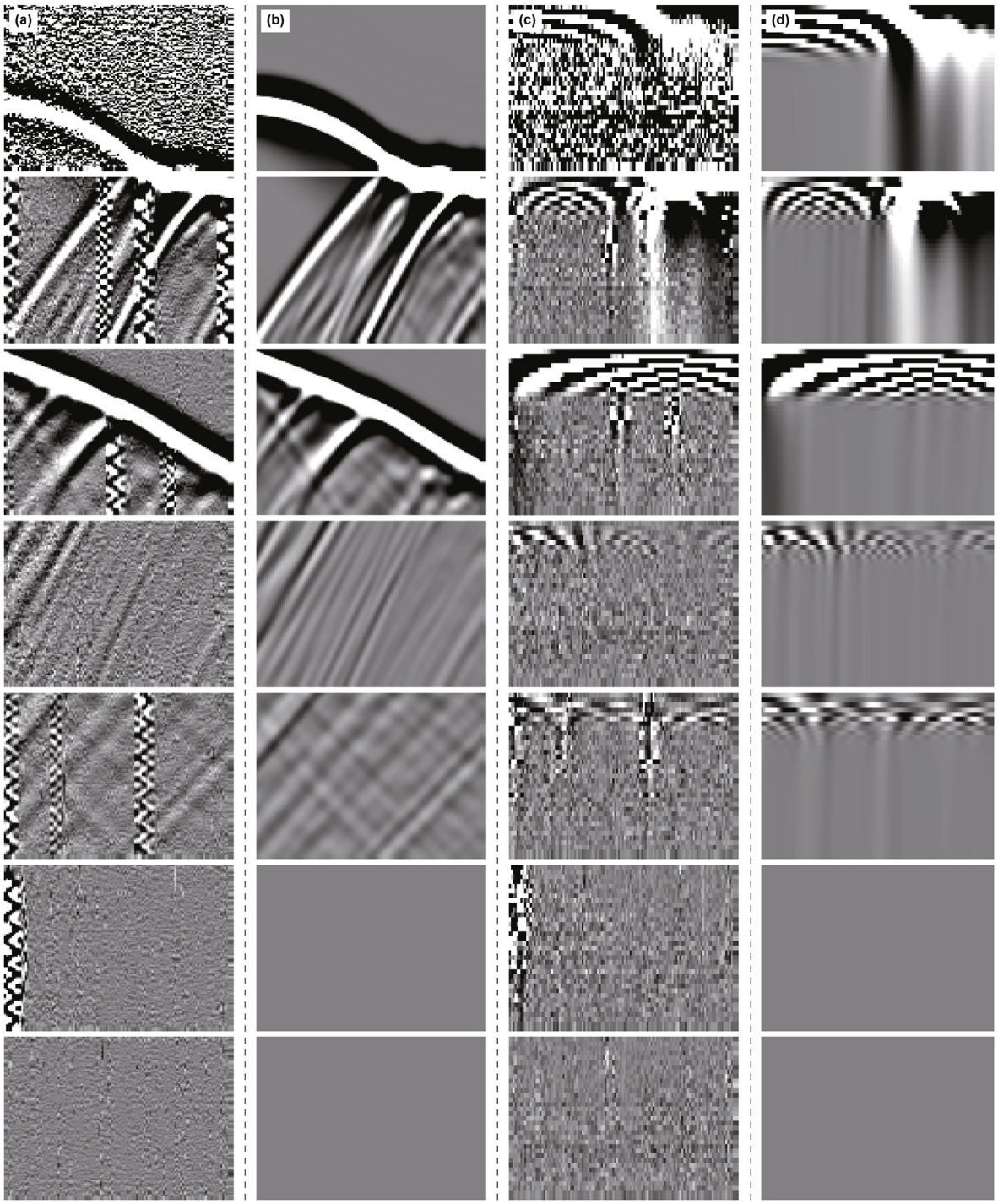

Fig. 10. Parts of the training datasets. **(a)** The noisy data in the time domain. **(b)** The clean data in the time domain. **(c)**–**(d)** The corresponding frequency domain data, respectively.

adaptive learning of noise characteristics and accurate noise extraction via EDCC-EMD, suppress these noise effectively.

To provide a clearer comparison, we offer zoomed-in views of the denoised data, marked by blue and red rectangles, in Figs. 14 and 15, respectively. Fig. 14(b) showcases the complex noisy data within the blue rectangle, characterized by random noise and three sets of strong coupling noise. This complexity results in a notably low SNR of −10.5 dB. When we compare the denoising results of these four methods, shown in Fig. 14(c)–(f), U-Net-*f* emerges as the most effective, enhancing the SNR of this noisy data from −10.5 dB to 10.9 dB, which is much better than the other three methods. Moreover, another example of noisy data, with an even lower SNR of −17.6 dB, is highlighted in the red rectangle and is presented in Fig. 15. This example is particularly challenging as the weak signals are completely buried in the noise. Remarkably, U-Net-*f* not only recovers the main structures of the weak signals but also suppresses the surrounding noise effectively. The zoomed-in figures clearly demonstrate that U-Net-*f* outperforms the other methods in removing random noise, and coupling noise while preserving weak signals.

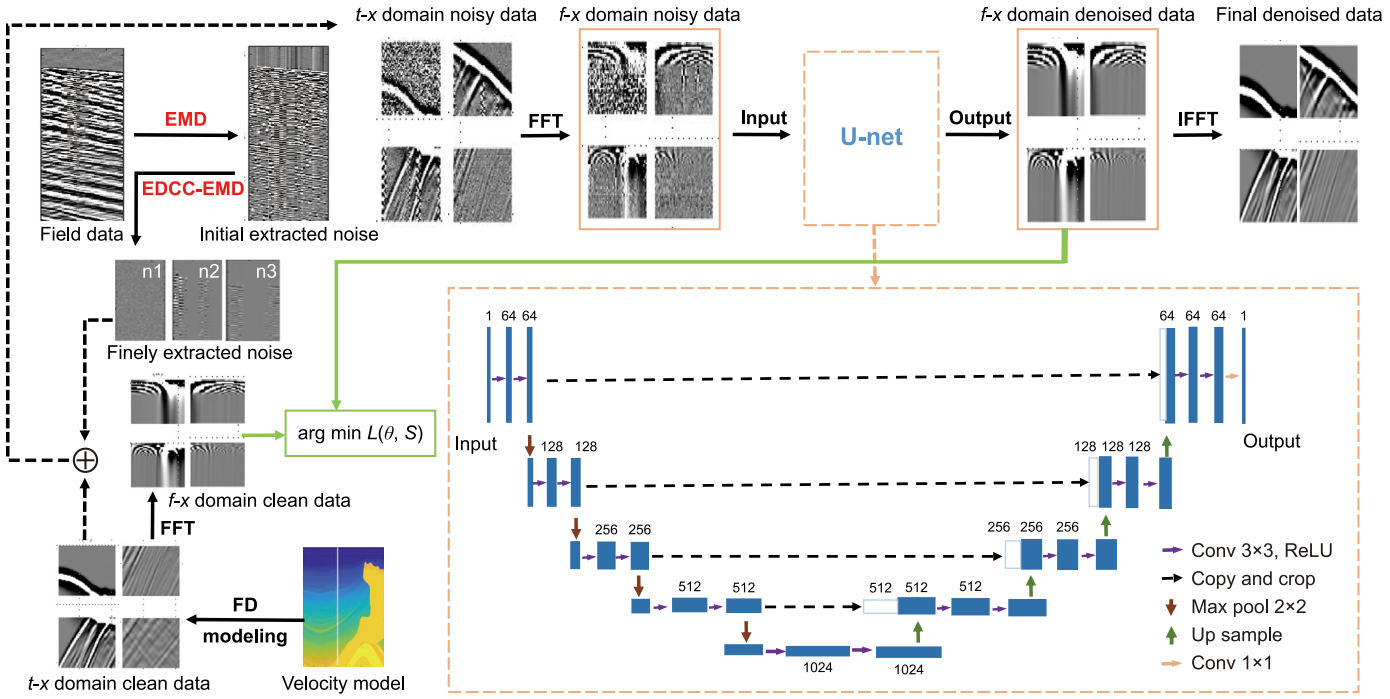

Fig. 11. Architecture of the network for DAS VSP denoising.

Table 1
Training configuration setting.

Parameters	Value
Convolution kernel size	3 × 3
Batch size	36
Learning rate	0.0001
Optimizer	Adam
Loss function	MSE
Epochs	50

Fourth, we focus on comparing the SNRs, structural similarity (SSIM) of the denoised results and the computational cost of the four methods, as detailed in Table 2. The proposed method, U-Net-*f*, significantly outperforms the others in terms of SNR improvement. Specifically, the SNR of the denoised data using U-Net-*f* is 22.6 dB, markedly higher than the results achieved by the other three methods. This indicates an impressive increase of 15 dB in SNR from the original noisy data. Concurrently, the U-Net-*f* achieves a significantly higher SSIM in its denoising outcomes compared to the other three methods. Also, this comparison highlights the efficiency of DL methods. For the synthetic data with the size of 3000 × 360 , the two U-Net methods take almost the same amount of time, less than 0.5 s, while the other two traditional methods take more than 1 s and 5 s respectively. Although the training time for the U-Net-*f* is more than 60 h, the denoising results validate its strong generalization ability, which will enable the U-Net-*f* to process new data without retraining. Therefore, the well trained U-net-*f* has higher efficiency than traditional methods in dealing with massive data. This efficiency is crucial in practical applications where processing speed and resource utilization are key considerations.

Fifth, we delve into the amplitude and phase details of the denoised products from U-Net-*f*, as illustrated in Fig. 16. The denoised trace is from Fig. 13(d), located at the near offset. Here, the denoised data (blue curve) aligns closely with the clean data (red curve) in both amplitude and phase (see Fig. 16(a)). This close match underscores the method's precision in noise removal. In

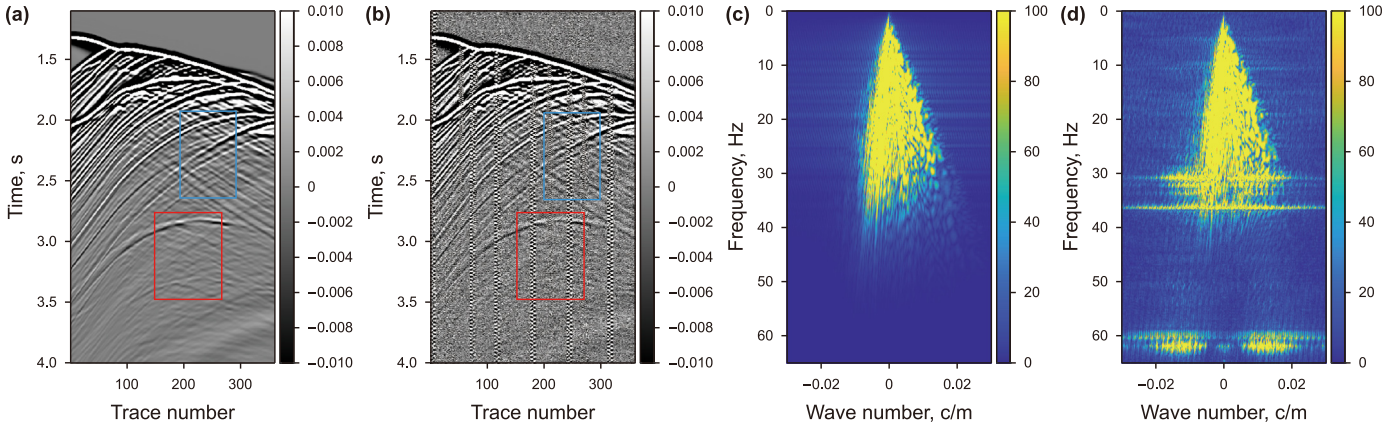

Fig. 12. The synthetic DAS VSP data to be tested. **(a)** The clean data. **(b)** The noisy data (7.6 dB). **(c)**–**(d)** The FK spectrum of the clean data and noisy data, respectively.

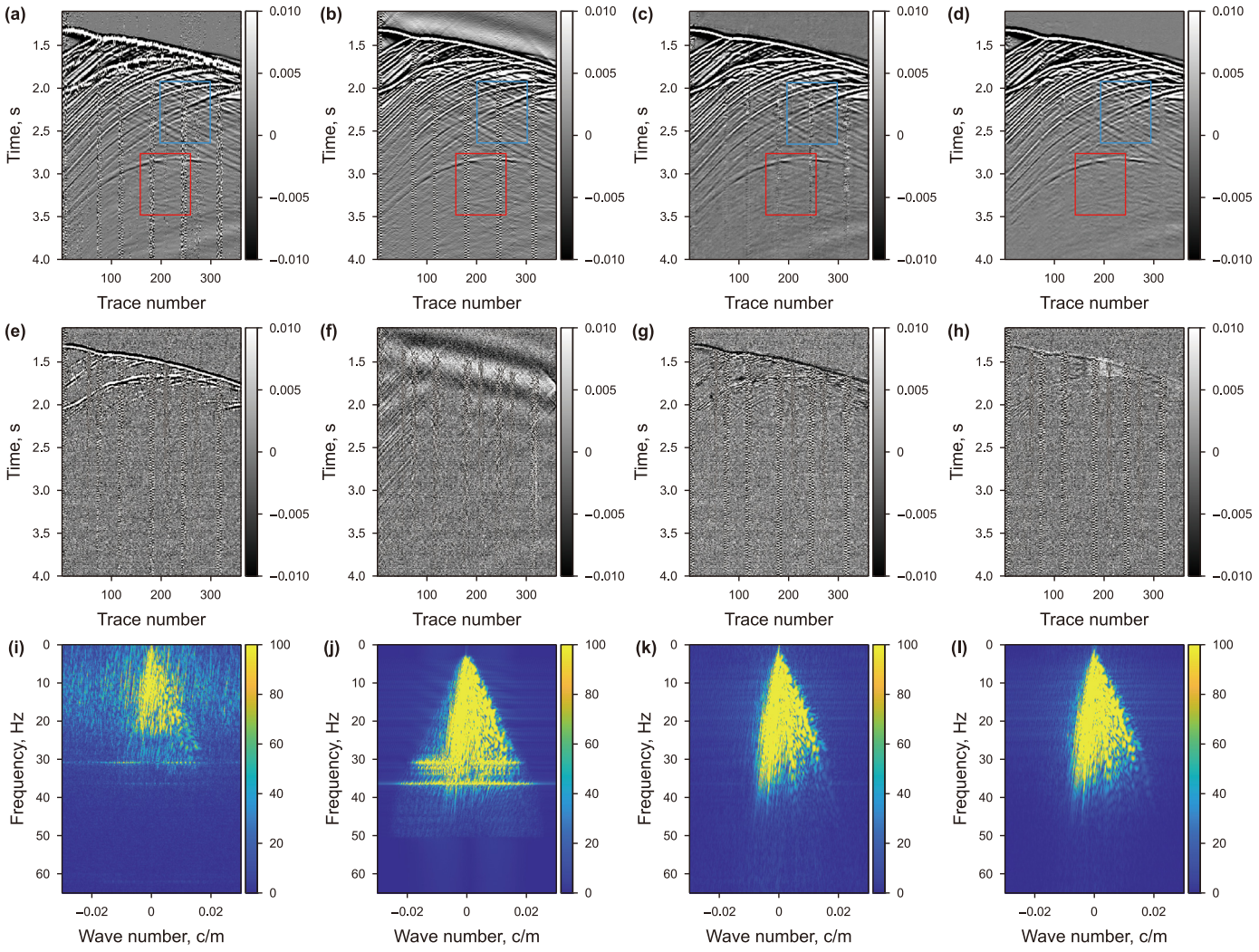

Fig. 13. Comparisons for the denoising results of synthetic data. Denoised data: **(a)** EMD; **(b)** FK; **(c)** U-Net-*t*; **(d)** U-Net-*f*. Removed noise: **(e)** EMD; **(f)** FK; **(g)** U-Net-*t*; **(h)** U-Net-*f*. FK spectrum of the denoised data: **(i)** EMD; **(j)** FK; **(k)** U-Net-*t*; **(l)** U-Net-*f*.

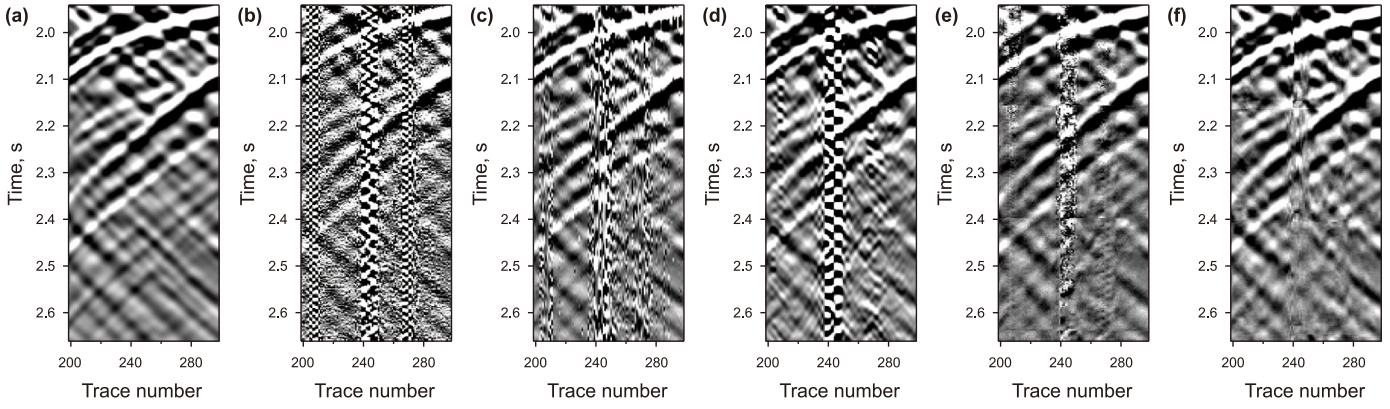

Fig. 14. Zoomed-view of the blue rectangles in the denoised products of Fig. 13. **(a)** The clean data. **(b)** The noisy data (−10.5 dB). **(c)** EMD (2.2 dB). **(d)** FK (0.7 dB). **(e)** U-Net-*t* (6.5 dB). **(f)** U-Net-*f* (10.9 dB).

areas where the removed noise (Fig. 16(b), pink curve) overlaps the pure noise (green curve), we observe an absence of signals, which suggests that U-Net-*f* effectively suppresses noise without damaging the underlying signal. Meanwhile, at the location marked by the red ellipse in Fig. 16(a), we note that our method corrects the phase distortions caused by noise, further indicating its denoising capability. The relative square error $(\sum (d-\hat{d})^2/\sum d^2 \times 100\%)$ between the denoised data $\hat{d}$ and the clean data d is 2.7%, which is an acceptable error range. The time-frequency spectrum of the

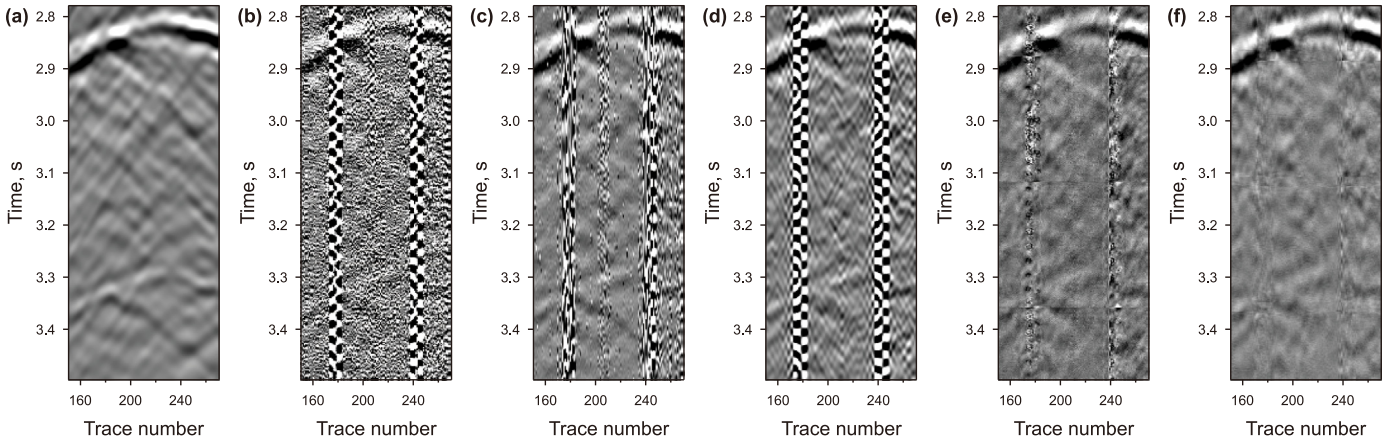

Fig. 15. Zoomed-view of the red rectangles in the denoised products of Fig. 13. **(a)** The clean data. **(b)** The noisy data (−17.6 dB). **(c)** EMD (0.2 dB). **(d)** FK (0.1 dB). **(e)** U-Net-*t* (2.1 dB). **(f)** U-Net-*f* (5.3 dB).

Table 2
SNRs and cost of the four methods.

Methods	SNR, dB	SSIM	Denoising cost, s	Training cost, h
Noisy data	7.6	0.8549	\	\
EMD	8.4	0.8851	1.6	\
FK	10.2	0.8907	5.3	\
U-Net-*t*	19.1	0.9761	0.3	31.2
U-Net-*f*	**22.6**	**0.9858**	**0.4**	62.3

clean data, the noisy data, the denoised data, and the removed noise (presented in Fig. 16(c)–(f)) verifies the effectiveness of the proposed method.

The comprehensive and detailed analyses of the above 5 aspects fully demonstrate the effectiveness of the proposed method in synthetic DAS VSP data denoising. Notably, our method achieves high-quality noise reduction without compromising the integrity of weak signals.

3.2. Field data

In realm of DL-based seismic denoising, the generalization ability is an important component in measuring the effectiveness of a denoising network. Our proposed method is now applied to a new field DAS VSP data, distinct from the field data used to extract noise labels. These two field datasets were acquired from two different survey areas. A comparison between the field data (Fig. 17(a)) and the synthetic data (Fig. 12(b)) highlights significant differences. The field DAS VSP data is more complex and has a lower SNR. The signals are contaminated by the strong random noise and coupling noise, posing a substantial challenge for denoising. The FK spectra of the field data reveal the complexity of this challenge. The coupling noise is spread across numerous narrow frequency bands, and what's worse is that most of these frequency bands are mixed with the signal's frequency bands.

We also compare and analyze the performance of these four methods from four aspects: denoised data, removed noise, FK spectrum of denoised data, and the details, as presented in Figs. 18 and 19. In terms of suppressing random noise, both FK and U-Net-*f* exhibit superior performance compared to EMD and U-Net-*t*, which is evident in the denoised data (Fig. 18(a)–(d)) and their corresponding FK spectrum (Fig. 18(i)–(l)). Regarding the attenuation of the coupling noise, the results are quite distinct. The denoised data from EMD, FK, and U-Net-*t*, shown in Fig. 18(a)–(c), retain coupling noise, particularly those noises in the near offset and those sharing the frequency bands with the signals. On the contrary, U-Net-*f* demonstrates its robust capability by simultaneously suppressing both random noise and coupling noise, as shown in Fig. 18(d).

For signal protection, the first three methods all damage the signals. This adverse effect is noticeable in the removed noise profiles (Fig. 18(e)–(g)), where significant signal residues are visible. However, our proposed method, U-Net-*f*, removes noise with almost no leaked signal (Fig. 18(h)), highlighting its effectiveness in preserving the signal while denoising. Fig. 18(i)–(l) display the FK spectrum corresponding to each denoising method. Notably, the spectrum processed by U-Net-*f* shows the least residual noise, significantly outperforming the other three methods. This impressive result showcases the exceptional learning ability of U-Net-*f*, which learns the key characteristics of random noise and coupling noise from the training data and can effectively separate them from the data without harming the signal. This performance of U-Net-*f* demonstrates that the training data we prepared is valid, and also, indicates that utilizing frequency domain data for training U-Net is an optimal strategy.

Meanwhile, we illustrate the zoomed-in views of specific areas marked by the red rectangles in Figs. 17 and 18. The data marked by the red rectangle in Fig. 17 are severely polluted by both random noise and coupling noise, as shown in Fig. 19 (a). We examine the denoising performance in more detail in the zoomed-in views to compare the denoising methods, as shown in Fig. 19(b)–(e). The magnified denoised data (Fig. 19 (e)) demonstrate that the U-Net-*f* can attenuate random noise and strong coupling noise simultaneously and enhance the identification of weak signals. We can see that in contrast to U-Net-*f*, the EMD method, limited by its theoretical characteristics, fails to suppress all of the noise and cannot preserve amplitude (Fig. 19 (b)). Similarly, the FK method struggles against coupling noise that shares the frequency with the signal (Fig. 19 (c)). For the U-Net-*t*, a small amount of coupling noise remains in the denoised data (Fig. 19 (d)).

Moreover, we quantified the denoising performance obtained from the field data using the S/N (Dash and Ahmed, 1970), expressed as follows:

$$S/N = 10\lg\frac{M\sum_{i=1}^{M-1}Q_{i,i+1}}{(M-1)\sum_{i=1}^{M}Q_{i,i}(0) - M\sum_{i=1}^{M-1}Q_{i,i+1}}, \tag{10}$$

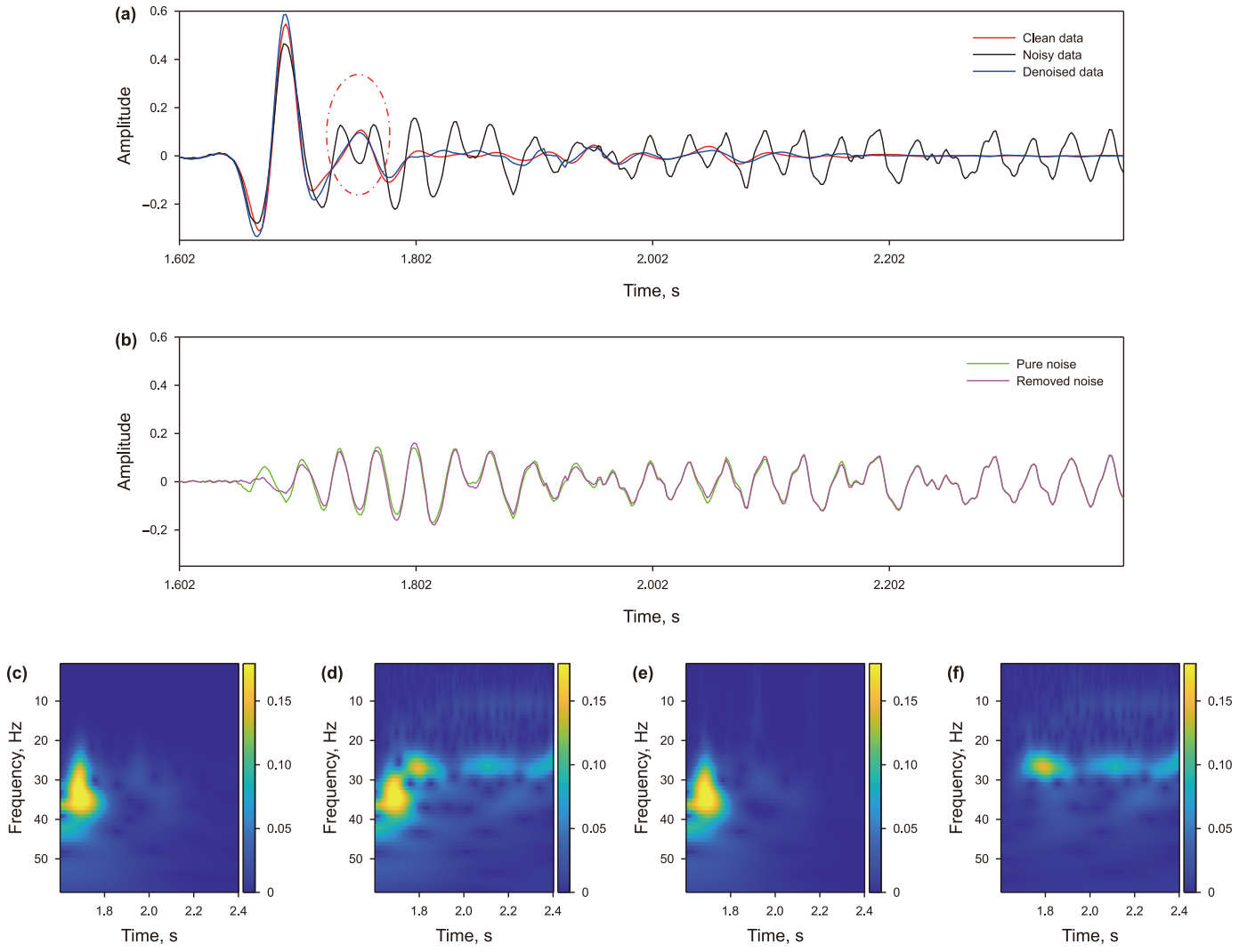

Fig. 16. Details of the amplitude and phase of data denoised by U-Net-*f*. **(a)** The third trace from Fig. 12(a) and (b), and Fig. 13(d), respectively. **(b)** The pure noise and the removed noise. Time-frequency spectrum: **(c)** clean data, **(d)** noisy data, **(e)** denoised data, **(f)** removed noise.

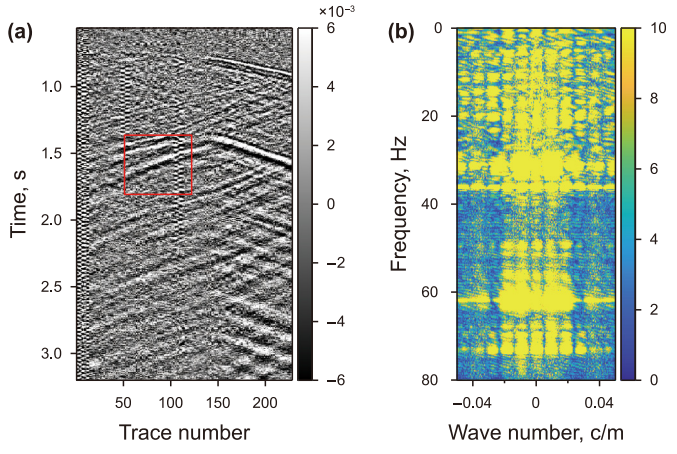

Fig. 17. The field DAS VSP data to be tested. **(a)** The noisy data. **(b)** The FK spectrum.

where $Q_{i,i+1}$ is the maximum value of the cross-correlation function between the two adjacent traces in the red rectangle, $Q_{i,i}$ is the maximum value of the autocorrelation function of the trace in the red rectangle, as shown in Fig. 18, and M is the trace number. The EMD method provides an S/N value of 4.58 dB, whereas the FK, U-Net-*t*, and the proposed U-Net-*f* methods produce an S/N of 13.85 dB, 14.01 dB, and 14.13 dB, respectively. Like synthetic data, the S/R of the denoised data obtained by the proposed U-Net-*f* is also the highest in field data testing. The U-Net-*f*, tailored for DAS VSP data, leverages the distinct characteristics of noise and signal in the frequency domain, which it more readily learns and distinguishes. As a result, U-Net-*f* achieves superior denoising results compared to the other three methods. Through the comprehensive analysis of the denoising results on the field DAS VSP data, the effectiveness of the proposed method is further verified.

4. Discussion

Some limitations and extensions of the proposed method are discussed. (1) The noise in the field DAS VSP data in this paper is mainly random noise and coupling noise, which is not completely consistent with the DAS noise in other research (Oboué et al., 2024). If the model trained in this paper is to be used to suppress other types of DAS noise, transfer learning is required. (2) The extracted noise in the training data set only comes from one optical fiber. If

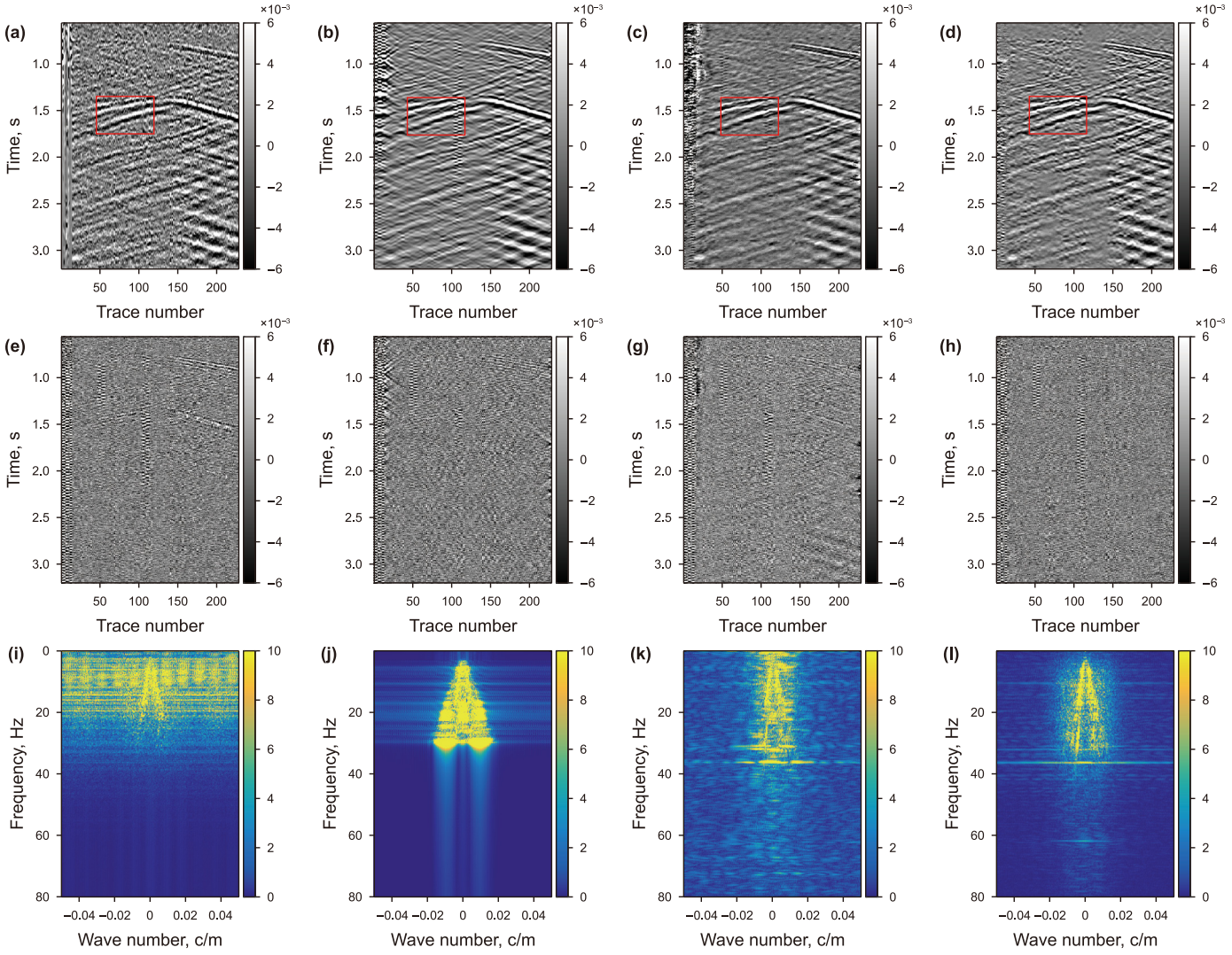

Fig. 18. Comparisons for denoising results of field DAS VSP data. Denoised data: **(a)** EMD; **(b)** FK; **(c)** U-Net-*t*; **(d)** U-Net-*f*. Removed noise: **(e)** EMD, **(f)** FK, **(g)** U-Net-*t*, **(h)** U-Net-*f*. FK spectra of the denoised data: **(i)** EMD, **(j)** FK, **(k)** U-Net-*t*, **(l)** U-Net-*f*.

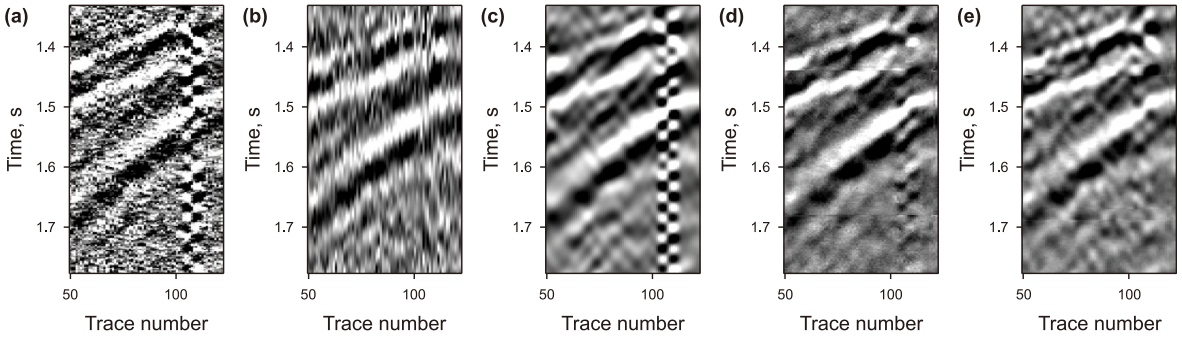

Fig. 19. Zoomed-view of the red rectangles in Figs. 17 and 18. **(a)** The noisy data. **(b)** EMD. **(c)** FK. **(d)** U-Net-*t*. **(e)** U-Net-*f*.

we have more fibers, more data with various coupling noises will be obtained, which will further improve the denoising result. (3) There are some signs of patch segment in the zoomed-in view of the denoised data, such as Fig. 14(e) and (f), Fig. 15(e) and (f), and Fig. 19(d) and (e). This phenomenon is caused by the fact that the network is trained and predicted in patches, which can be solved by improving the computing power for larger patch-size training and prediction. (4) In our study, the network trained in the frequency-spatial domain performs effectively in noise suppression, but this does not mean that other domains such as the singular spectrum (Feng and Li, 2021), frequency-time domain (Iqbal, 2022), and the DWT domain (Wang et al., 2023a) are not applicable. It needs to be compared through a large number of experiments, but this is not the focus of this paper. We aim to enhance the understanding and

analysis of DAS VSP data by deep learning methods in the frequency-spatial domain, thereby expanding the depth and breadth of seismology. This rigorous analytical approach enables accurate identification of subtle DAS VSP signals, mitigates noise interference in data interpretation, and establishes a more robust foundation for seismology and related fields. Our future research will involve comparing various transforms and assessing the influence of frequency-spatial domain information on other seismic tasks.

5. Conclusion

To suppress the random noise and coupling noise in distributed acoustic sensing (DAS) vertical seismic profile (VSP), we designed an exponential decay curve-constrained empirical mode decomposition (EDCC-EMD) analysis-based supervised denoising network. This approach effectively overcomes a primary limitation of supervised denoising networks, that is, the difficulty in obtaining accurate labels for training in real-world scenarios. Through the EDCC-EMD method, we can swiftly and accurately extract pure random noise and coupling noise from field DAS VSP data. This extracted noise forms a robust training dataset, enabling the network to precisely learn the distinct noise characteristics inherent in DAS VSP data, thereby facilitating their effective separation from the signal. Our findings demonstrate that the EDCC-EMD method effectively addresses these two types of noises with distinct properties. This affirms that the EDCC-EMD can be applied not only to DAS VSP data but also more broadly across fields such as seismic data collected by conventional geophones. Meanwhile, our methodology incorporated frequency-domain data for training the network, deviating from the conventional utilization on time-domain data. This shift to frequency-domain data is strategic, capitalizing on the clearer distinction of noise and signal characteristics in this domain, which enhances the network's learning efficiency and effectiveness in noise-signal separation.

The synthetic and field examples demonstrated the exceptional denoising capability of the proposed method. With synthetic data, our method, particularly the U-Net-*f* model, significantly improved the signal-to-noise ratio by 15 dB. More impressively, in field data, our method achieved remarkable denoising products even in severely noise-polluted DAS VSP scenarios, highlighting its robust generalization ability. When compared with other methods, including EMD, FK, and U-Net-*t*, our method consistently outperforms in terms of noise attenuation and signal preservation, showcasing its superiority in handling DAS VSP data.

CRediT authorship contribution statement

Huan-Huan Tang: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Shi-Jun Cheng:** Writing – review & editing. **Wu-Qun Li:** Data curation. **Wei-Jian Mao:** Writing – review & editing, Investigation, Funding acquisition.

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.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Nos. 42404140, 42130808) and the National Key Research and Development Program of China under grant 2021YFA0716802. We thank Professor Xiang-Fang Zeng from Innovation Academy for Precision measurement Science and Technology, Chinese Academy of Sciences for his valuable discussions.

References

Alberto, D.C.F.C., Meles, G.A., Curtis, A., 2017. Elastic internal multiple analysis and attenuation using Marchenko and interferometric methods. Geophysics 82 (2), 1–12. https://doi.org/10.1190/geo2016-0162.1.

Anderson, R.G., McMechan, G.A., 1988. Noise-adaptive filtering of seismic shot records. Geophysics 53 (5), 638–649. https://doi.org/10.1190/1.1442498.

Anvari, R., Mohammadi, M., Kahoo, A.R., et al., 2020. Random noise attenuation of 2D seismic data based on sparse low-rank estimation of the seismic signal. Comput. Geosci. 135 (C), 1–12. https://doi.org/10.1016/j.cageo.2019.104376.

Alsdorf, D., 1997. Noise reduction in seismic data using Fourier correction coefficient filtering. Geophysics 62 (5), 1617–1627. https://doi.org/10.1190/1.1444264.

Battista, B.M., Knapp, C., McGee, T., et al., 2007. Application of the empirical mode decomposition and Hilbert-Huang transform to seismic reflection data. Geophysics 72 (2), 29–37. https://doi.org/10.1190/1.2437700.

Bekara, M., Van der Baan, M., 2009. Random and coherent noise attenuation by empirical mode decomposition. Geophysics 74 (5), 89–98. https://doi.org/10.1190/1.3157244.

Binder, G., Titov, A., Liu, Y., et al., 2020. Modeling the seismic response of individual hydraulic fracturing stages observed in a time-lapse distributed acoustic sensing vertical seismic profiling survey. Geophysics 85 (4), 225–235. https://doi.org/10.1190/geo2019-0819.1.

Birnie, C., Ravasi, M., Liu, S., et al., 2021. The potential of self-supervised networks for random noise suppression in seismic data. Artif. Intell. Geosci. 2 (1), 47–59. https://doi.org/10.1016/j.aiig.2021.11.001.

Cai, Z., Cheng, T.H., Lu, C., et al., 2001. Efficient wavelet-based image denoising algorithm. Electron. Lett. 37 (11), 683–685. https://doi.org/10.1049/el:20010466.

Chen, J., Ning, J., Chen, W., et al., 2019. Distributed acoustic sensing coupling noise removal based on sparse optimization. Interpretation 7 (2), 373–382. https://doi.org/10.1190/INT-2018-0080.1.

Chen, J., Chen, J., Chao, H., et al., 2018. Image Blind denoising with generative adversarial network based noise modeling. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/CVPR.2018.00333.

Chen, Y., Zhou, C., Yuan, J., et al., 2014. Application of empirical mode decomposition to random noise attenuation of seismic data. J. Seismic Explor. 23 (5), 481–495.

Chen, Y., Zhou, Y., Chen, W., et al., 2017. Empirical low-rank approximation for seismic noise attenuation. IEEE Trans. Geosci. Rem. Sens. 55 (8), 4696–4711. https://doi.org/10.1109/TGRS.2017.2698342.

Cheng, S., Harsuko, R., Alkhalifah, T., 2024a. Meta-Processing: a robust framework for multi-tasks seismic processing. Surv. Geophys. 451, 1081–1116. https://doi.org/10.1007/s10712-024-09837-9.

Cheng, S., Cheng, Z., Jiang, C., et al., 2024b. An effective self-supervised learning method for various seismic noise attenuation. Geophysics 89 (6), 589–604. https://doi.org/10.1190/geo2023-0656.1.

Dash, B.P., Ahmed, O.K., 1970. Determination of signal and noise statistics using correlation theory. Geophysics 35 (1), 24–32. https://doi.org/10.1190/1.1440077.

Dong, X., Li, Y., Zhong, T., et al., 2020. Random and coherent noise suppression in DAS-VSP data by using a supervised deep learning method. Geosci. Rem. Sens. Lett. IEEE 19, 1–5. https://doi.org/10.1109/LGRS.2020.3023706.

Dragoset, B., Verschuur, E., Moore, I., et al., 2010. A perspective on 3D surface-related multiple elimination. Geophysics 75 (5), 245–261. https://doi.org/10.1190/1.3475413.

Falk, T., Mai, D., Bensch, R., et al., 2019. Author Correction: U-Net: deep learning for cell counting, detection, and morphometry. Nat. Methods 16 (4), 351. https://doi.org/10.1038/s41592-019-0356-4.

Feng, Q., Li, Y., 2021. Denoising deep learning network based on singular spectrum analysis—DAS seismic data denoising with multichannel SVDDCNN. IEEE Trans. Geosci. Rem. Sens. 60, 1–11. https://doi.org/10.1109/TGRS.2021.3071189.

Feng, Q., Wang, S., Li, Y., 2024. Analysis of DAS seismic noise generation and elimination process based on mean-SDE diffusion model. IEEE Trans. Geosci. Rem. Sens. 62, 1–13. https://doi.org/10.1109/TGRS.2024.3362853.

Foster, D.J., Mosher, C.C., 1992. Suppression of multiple reflections using the Radon transform. Geophysics 57 (3), 386–395. https://doi.org/10.1190/1.1443253.

Gan, S., Wang, S., Chen, Y., Qu, S., Zu, S., 2016. Velocity analysis of simultaneous-source data using high-resolution semblance-Coping with the strong noise. Geophys. J. Int. 204 (2), 768–779. https://doi.org/10.1016/j.cageo.2023.105464.

Geetha, K., Hota, M.K., 2023. Seismic random noise attenuation using optimal empirical wavelet transform with a new wavelet thresholding technique. IEEE Sens. J. 24 (1), 596–606. https://doi.org/10.1109/JSEN.2023.3334819.

Gómez, J.L., Velis, D.R., 2016. A simple method inspired by empirical mode decomposition for denoising seismic data. Geophysics 81 (6), 403–413. https://doi.org/10.1190/geo2015-0566.1.

Huang, N.E., Shen, Z., Long, S.R., et al., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. Roy. Soc. A 454 (1971), 903–995. https://doi.org/10.1098/rspa.1998.0193.

Iqbal, N., 2022. DeepSeg: deep segmental denoising neural network for seismic data. IEEE Transact. Neural Networks Learn. Syst. 34 (7), 3397–3404. https://doi.org/10.1109/TNNLS.2022.3205421.

Jain, V., Seung, S., 2008. Natural image denoising with convolutional networks. In: Proceedings of the 22nd International Conference on Neural Information Processing Systems, pp. 769–776.
Jin, Y., Wu, X., Chen, J., et al., 2018. Seismic data denoising by deep-residual networks. In: 88th Annual International Meeting. SEG Expanded Abstracts, pp. 4593–4597. https://doi.org/10.1190/segam2018-2998619.1.
Kabir, M.M.N., Verschuur, D.J., 1995. Restoration of missing offsets by parabolic Radon transform. Geophys. Prospect. 43 (3), 347–368. https://doi.org/10.1111/j.1365-2478.1995.tb00257.x.
Kostov, C., De Melo, F.X., Raj, A., et al., 2015. Multiple attenuation for shallow-water surveys: notes on old challenges and new opportunities. Lead. Edge 34 (7), 760–768. https://doi.org/10.1190/tle34070760.1.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539:521.
Liu, C., Liu, Y., Yang, B., et al., 2006. A 2D multistage median filter to reduce random seismic noise. Geophysics 71 (5), 105–110. https://doi.org/10.1190/1.2236003.
Li, Q., Vasudevan, K., Cook, F.A., 1997. 3-D coherency filtering. Geophysics 62 (4), 1310–1314. https://doi.org/10.1190/1.1444232.
Li, S., Liu, B., Ren, Y., et al., 2019. Deep-learning inversion of seismic data. IEEE Trans. Geosci. Rem. Sens. 58 (3), 2135–2149. https://doi.org/10.1109/TGRS.2019.2953473.
Liu, S., Cheng, S., Alkhalifah, T., 2024. Gabor-based learnable sparse representation for self-supervised denoising. IEEE Trans. Geosci. Rem. Sens. 62, 1–16. https://doi.org/10.1109/TGRS.2024.3353315.
Liu, T.T., Chen, Y., 2016. Random noise attenuation based on EMD and MSSA in fx domain. Geophys. Prospect. Pet. 55 (1), 67–75. https://doi.org/10.3969/j.issn.1000-1441.2016.01.009.
Miller, D., Parker, T., Kashikar, S., et al., 2012. Vertical seismic profiling using a fibre-optic cable as a distributed acoustic sensor. In: 74th EAGE Conference and Exhibition Incorporating EUROPEC 2012, p. 293. https://doi.org/10.3997/2214-4609.20148799.
Mandelli, S., Lipari, V., Bestagini, P., et al., 2019. Interpolation and Denoising of Seismic Data Using Convolutional Neural Networks arxiv preprint arxiv: 1901.07927.
Mateeva, A., Lopez, J., Potters, H., et al., 2014. Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling. Geophys. Prospect. 62 (4), 679–692. https://doi.org/10.1111/1365-2478.12116.
Oboué, Y.A.S.I., Wei, C., Saad, O.M., et al., 2023. Adaptive damped rank-reduction method for random noise attenuation of three-dimensional seismic data. Surv. Geophys. 44, 847–875. https://doi.org/10.1007/s10712-022-09756-7.
Oboué, Y.A.S.I., Chen, Y., Fomel, S., et al., 2024. An advanced median filter for improving the signal-to-noise ratio of seismological datasets. Comput. Geosci. 182, 1–24. https://doi.org/10.1016/j.cageo.2023.105464.
Qian, F., Hua, H., Wen, Y., et al., 2024. Unsupervised intense VSP coupling noise suppression with iterative robust deep learning. IEEE Trans. Geosci. Rem. Sens. 62, 1–17. https://doi.org/10.1109/TGRS.2023.3349272.
Richardson, A., Feller, C., 2019. Seismic Data Denoising and Deblending Using Deep Learning arxiv preprint arxiv:1907.01497.
Ross, Z.E., Meier, M.A., Hauksson, E., 2018. P wave arrival picking and first-motion polarity determination with deep learning. J. Geophys. Res. 123 (6), 5120–5129. https://doi.org/10.1029/2017JB015251.
Saad, O.M., Chen, Y., 2020. Deep denoising autoencoder for seismic random noise attenuation. Geophysics 85 (4), 367–376. https://doi.org/10.1190/geo2019-0468.1.
Schimmel, M., Gallart, J., 2007. Frequency-dependent phase coherence for noise suppression in seismic array data. J. Geophys. Res. 112, B04303. https://doi.org/10.1029/2006JB004680.
Shao, J., Wang, Y., Liang, X., et al., 2022. Siamese network based noise elimination of artificial seismic data recorded by distributed fiber-optic acoustic sensing. Chin. J. Geophys. 65 (9), 3599–3609. https://doi.org/10.6038/cjg2022P0919.
Song, C., Guo, S., Xiong, C., et al., 2024. Regularized deep learning for unsupervised random noise attenuation in poststack seismic data. J. Geophys. Eng. 21 (1), 60–67. https://doi.org/10.1093/jge/gxad094.
Song, H., Mao, W., Tang, H., et al., 2022. Multiple attenuation based on connected-component analysis and high-resolution parabolic Radon transform. Journal of J. Appl. Geophys. 199. https://doi.org/10.1016/j.jappgeo.2022.104580.
Staring, M., Zhang, L., Thorbecke, J., et al., 2020. A new role for adaptive filters in Marchenko equation-based methods for the attenuation of internal multiples. In: 82nd EAGE Conference and Exhibition Incorporating EUROPEC 2020. https://doi.org/10.3997/2214-4609.202012132.
Stein, R.A., Bartley, N.R., 1983. Continuously time-variable recursive digital band-pass filters for seismic signal processing. Geophysics 48 (6), 702–712. https://doi.org/10.1190/1.1441500.
Taner, M.T., O'Doherty, R.F., Koehler, F., 1995. Long period multiple suppression by predictive deconvolution in the x–t domainl1. Geophys. Prospect. 43 (4), 433–468. https://doi.org/10.1111/j.1365-2478.1995.tb00261.x.
Thorson, J.R., Claerbout, J.F., 1985. Velocity-stack and slant-stack stochastic inversion. Geophysics 50 (12), 2727–2741. https://doi.org/10.1190/1.1441893.
Wang, K., Hu, T., Zhao, B., et al., 2023a. An unsupervised deep learning method for direct seismic deblending in shot domain. IEEE Trans. Geosci. Rem. Sens. 61, 1–12. https://doi.org/10.1109/TGRS.2023.3298054.
Wang, Y., Liu, X., Gao, F., et al., 2020. Robust vector median filtering with a structure- adaptive implementation. Geophysics 85 (5), 407–414. https://doi.org/10.1190/geo2020-0012.1.
Wang, S., Song, P., Tan, J., et al., 2023b. Deblending of seismic data in the wavelet domain via a convolutional neural network based on data augmentation. Geophys. Prospect. 72 (1), 213–228. https://doi.org/10.1111/1365-2478.13277.
Wang, X., Sui, Y., Wang, W., et al., 2023c. Random noise attenuation by self-supervised learning from single seismic data. Math. Geosci. 55 (3), 401–422. https://doi.org/10.1007/s11004-022-10032-y.
Wilson, C., Guitton, A., 2007. Teleseismic wavefield interpolation and signal extraction using high-resolution linear radon transforms. Geophys. J. Int. 168 (1), 171–181. https://doi.org/10.1111/j.1365-246X.2006.03163.x.
Xu, Y., Zhu, H., Cao, S., et al., 2024. DAS coupling noise suppression based on MCA–FK. Acta Geophys. 72, 2465–2474. https://doi.org/10.1007/s11600-023-01225-y.
Xue, Y., Man, M., Zu, S., et al., 2017. Amplitude-preserving iterative deblending of simultaneous source seismic data using high-order Radon transform. J. Appl. Geophys. 139, 79–90. https://doi.org/10.1016/j.jappgeo.2017.02.010.
Yang, L., Chen, W., Liu, W., et al., 2020. Random noise attenuation based on residual convolutional neural network in seismic datasets. IEEE Access 8, 30271–30286. https://doi.org/10.1109/ACCESS.2020.2972464.
Yang, L., Fomel, S., Wang, S., et al., 2023. Denoising distributed acoustic sensing data using unsupervised deep learning. Geophysics 88 (4), 317–332. https://doi.org/10.1190/geo2022-0460.1.
Yin, H., Cao, J., Yang, H., et al., 2024. An adaptive seismic random noise attenuation method based on Engl criterion using curvelet transform. J. Appl. Geophys. 227. https://doi.org/10.1016/j.jappgeo.2024.105416.
Yu, G., Cai, Z., Chen, Y., et al., 2016. Walkaway VSP using multimode optical fibers in a hybrid wireline. Lead. Edge 35 (7), 615–619. https://doi.org/10.1190/tle35070615.1.
Yu, S., Ma, J., Wang, W., 2019. Deep learning for denoising. Geophysics 84 (6), 333–350. https://doi.org/10.1190/geo2018-0668.1.
Yuan, Y., Li, Y., Li, G., et al., 2022. Adaptive ground-roll attenuation using local nonlinear filtering. Comput. Geosci. 164, 105124. https://doi.org/10.1016/j.cageo.2022.105124.
Zhao, Y., Li, Y., Dong, X., et al., 2018. Low-frequency noise suppression method based on improved DnCNN in desert seismic data. Geosci. Rem. Sens. Lett. IEEE 16 (5), 811–815. https://doi.org/10.1109/TGRS.2019.2938836.
Zhang, H., Yang, H., Li, H., et al., 2018. Random noise attenuation of non-uniformly sampled 3D seismic data along two spatial coordinates using non-equispaced curvelet transform. J. Appl. Geophys. 151, 221–233. https://doi.org/10.1016/j.jappgeo.2018.02.018.
Zhang, H., Yang, X., Ma, J., 2020. Can learning from natural image denoising be used for seismic data interpolation? Geophysics 85 (4), 115–136. https://doi.org/10.1190/geo2019-0243.1.
Zhang, K., Zuo, W., Chen, Y., et al., 2017. Beyond a Gaussian denoiser: residual learning of deep cnn for image denoising. IEEE Trans. Image Process. 26 (7), 3142–3155. https://doi.org/10.1109/TIP.2017.2662206.
Zhang, M., Liu, Y., Bai, M., et al., 2019. Seismic noise attenuation using unsupervised sparse feature learning. IEEE Trans. Geosci. Rem. Sens. 57 (12), 9709–9723. https://doi.org/10.1109/TGRS.2019.2928715.
Zhang, Y., Wang, B., 2023. Unsupervised seismic random noise attenuation by a recursive deep image prior. Geophysics 88 (6), 473–485. https://doi.org/10.1190/geo2022-0612.1.