



Self-supervised Multistep Seismic Data Deblending

Xinyi Chen¹ · Benfeng Wang¹

Received: 11 December 2022 / Accepted: 8 July 2023 / Published online: 18 August 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

The potential of blended seismic acquisition to improve acquisition efficiency and cut acquisition costs is still open, particularly with efficient deblending algorithms to provide accurate deblended data for subsequent processing procedures. In recent years, deep learning algorithms, particularly supervised algorithms, have drawn much attention over conventional deblending algorithms due to their ability to nonlinearly characterize seismic data and achieve more accurate deblended results. Supervised algorithms require large amounts of labeled data for training, yet accurate labels are rarely accessible in field cases. We present a self-supervised multistep deblending framework that does not require clean labels and can characterize the decreasing blending noise level quantitatively in a flexible multistep manner. To achieve this, we leverage the coherence similarity of the common shot gathers (CSGs) and the common receiver gathers (CRGs) after pseudo-deblending. The CSGs are used to construct the training data adaptively, where the raw CSGs are regarded as the label with the corresponding artificially pseudo-deblended data as the initial training input. We employ different networks to quantitatively characterize decreasing blending noise levels in multiple steps for accurate deblending with the help of a blending noise estimation–subtraction strategy. The training of one network can be efficiently initialized by transfer learning from the optimized parameters of the previous network. The optimized parameters trained on CSGs are used to deblend all CRGs of the raw pseudo-deblended data in a multistep manner. Tests on synthetic and field data validate the proposed self-supervised multistep deblending algorithm, which outperforms the multilevel blending noise strategy.

Keywords Semi-supervised deblending · Coherence similarity · Multistep deblending · Multilevel blending noise strategy

Article Highlights

- Considering the coherence similarity of CSGs and CRGs after pseudo-deblending, we adaptively construct the training sets using the CSGs as the desired output and the constructed data with blending noise contamination as the initial training input

✉ Benfeng Wang
wbf1232007@126.com

¹ State Key Laboratory of Marine Geology, Tongji University, Shanghai 200092, China

- Multistep deblending that employs different individual networks with a flexibly designed loss function is used to evaluate the remaining blending noise, and the structural similarity (SSIM) is incorporated into the subsequent loss function to attenuate weak blending noise precisely
- Numerically blended synthetic and field data examples demonstrate the validity of the proposed self-supervised multistep deblending algorithm in accurately obtaining deblended data without clean label requirements

1 Introduction

Conventional land and marine seismic data acquisitions fire seismic sources with long time intervals to avoid signal interference, and the acquisition efficiency is low with high acquisition costs. Blended acquisition fires multiple shots almost simultaneously using a dithering strategy (Berkhout et al. 2008; Hampson et al. 2008) to significantly improve acquisition efficiency, especially for wide azimuth (Walker et al. 2013) and wide bandwidth (Poole et al. 2014) seismic data acquisition.

There are two categories of algorithms for blended data applications. One option is to directly perform reverse time migration (RTM) (Dai et al. 2011; Verschuur and Berkhout 2011) or full waveform inversion (FWI) (Zhang et al. 2018), which requires special processing algorithms to suppress artifacts caused by the blending noise during RTM and FWI procedures. For blended data with a high blending fold, however, the artifacts caused by the blending noise remain, contaminating the final RTM images and making subsequent seismic interpretation more challenging. The other option attenuates the blending noise first and then uses conventional, mature seismic processing workflows for the obtained deblended results (Mahdad 2012). This latter option has recently attracted much attention and is still open.

Deblending can be viewed as a denoising issue, and deblended data can be obtained by filtering-based algorithms. Although a receiver records the information from multiple blended sources continuously during blended acquisition, the signal of pseudo-deblended data is coherent in certain domains, like the common receiver domain (CRD), whereas the blending noise is randomized along spaces (Beasley 2008). In these domains, deblending becomes a standard denoising procedure that distinguishes the desired coherent signal from the randomized blending noise. Median filter (MF) and its variations are always used for deblending since MF is well-known for random, spike-like noise removal (Liu et al. 2009). Huo et al. (2012) developed a vector median filter from the conventional scalar version and used a multi-directional vector median filter for deblending in the common mid-point domain (CMD). Chen et al. (2014b) employed a recursive strategy for deblending in a closed-loop manner, which updated the deblended result using MF on normal-moveout corrected data in the CMD that required a new velocity estimation from the deblended result. Gan et al. (2016) proposed a structural-oriented median filter for deblending, which flattened the pseudo-deblended data in local spatial windows and then applied MF to attenuate the blending noise. The above-mentioned methods based on filtering are efficient; however, the accuracy is still open to improvement. From another different perspective, deblending can also be resolved as an ill-posed inversion problem, commonly with certain prior knowledge constraints, such as low-rank (Cheng and Sacchi, 2015b), sparsity (Kontakis and Verschuur 2014), or coherency (Huang et al. 2017). Sparse constraints are used for a variety of sparse transforms, such as the Fourier transform (Abma et al., 2010),

Radon transform (Ibrahim and Sacchi 2014), seislet transform (Chen et al. 2014a), and curvelet transform (Wang and Geng 2019), etc. Chen (2015) proposed an iterative deblending algorithm with multiple constraints using shaping regularization, combining iterative seislet thresholding with a local orthogonalization strategy. For field blended data deblending, Zu et al. (2017) analyzed different factors that can affect the final deblending performance, highlighting the importance of considering these factors during blended acquisition to avoid potential failures. Lin et al. (2021) implemented a robust iterative deblending algorithm by adopting a robust singular spectrum analysis (SSA) filter for each frequency slice. In cases where shot time is inaccurate, Chen et al. (2023) designed a joint inversion framework to simultaneously estimate the deblended data and the shot-time vector. The key to the above-mentioned filtering and inversion-based deblending algorithms is the selection of appropriate parameters, such as the filter length, types of sparse transforms and corresponding thresholds, or the designed rank. However, these parameters are typically chosen by trial and error according to individual experience and may vary from one dataset to another, requiring fine-tuning in practice. Besides, the computational costs of traditional deblending algorithms increase significantly for multidimensional and high-resolution blended seismic data.

Deep learning (DL) has been widely used in various geophysical issues, such as seismic data processing and interpretation (Huang et al. 2022; Park and Sacchi 2020; Yang et al. 2020). Among various applications, the denoising issue attracts lots of interest, especially when well-structured networks like U-net (Ronneberger et al. 2015) are used to capture high-level nonlinear relationships between the input (noise-contaminated data) and the desired output (clean data) using a large number of training sets. As for DL-based deblending issues, three main aspects have always been considered since the CNN algorithm was introduced into seismic deblending (Sun et al. 2020; Zu et al. 2020), i.e., independence or weak dependence on a large volume of clean labels (Baardman et al. 2020; Sun et al. 2022), higher deblending accuracy (Wang et al. 2023), and better generalization (Wang et al. 2021a). To avoid clean label requirements, unsupervised or self-supervised algorithms appear with physical constraints on blended data (Wang et al. 2022c; Xue et al. 2022). Wang and Hu (2022) designed a self-supervised deblending workflow in which the training procedure utilized common shot gathers (CSGs) and the test deblending step was implemented on common receiver gathers (CRGs) via the optimized network based on the coherence similarity between CSGs and CRGs. Wang et al. (2021b) designed a blind-trace network, inspired by the blind-spot network, for self-supervised deblending that directly mapped pseudo-deblended data to the coherent signal in the CRD. The residual between artificially blended data and the observed blended data was measured to construct the objective function. Besides, there are many attempts and enhancements for supervised deblending algorithms. Xu et al. (2022) integrated the inversion-based deblending algorithm with DL algorithm, employing a deep CNN network trained on pseudo-deblended CSGs as a Gaussian denoiser in the plug and play (PnP) deblending algorithm. To enhance deblending accuracy and generalization, Wang et al. (2021a) used transfer learning to fine-tune a previously optimized model with parts of labeled field data to adapt field data deblending. A multi-resolution ResUnet was also introduced (Wang et al. 2022a), with a multilevel blending noise strategy (MBNS) simulating the decreasing blending noise level during iterative deblending. Later, the MBNS was also extended to simultaneous seismic data deblending and recovery (Wang et al. 2022b) for incomplete blended data. However, it qualitatively mimics the decreasing blending noise level during iterative deblending using several fixed scalars predefined by trial and error, which deviates from the fact during iterative applications. Thus, a multistep deblending algorithm was designed to quantitatively

characterize the decreasing blending noise (Wang et al. 2023), further improving the deblending accuracy.

In this paper, we present a self-supervised multistep deblending workflow to quantitatively characterize nonlinear features of different levels of remaining blending noise during deblending procedures, without clean label requirements. Firstly, we extract the CSGs \mathbf{d}_{csg} of pseudo-deblended data as the desired output. Artificially pseudo-deblended data $\mathbf{d}_{csg_pdb} = \mathbf{\Gamma}^H \mathbf{\Gamma} \mathbf{d}_{csg}$ based on CSGs is regarded as the initial input to generate adaptive training pairs with a blending operator $\mathbf{\Gamma}$ and its adjoint operator. Secondly, we introduce the multistep deblending algorithms to quantify the decreasing blending noise levels with the help of the blending noise estimation–subtraction strategy. In the first deblending step, the network is trained to recover the dominant signal by attenuating the blending noise. With the signal estimated from the first deblending step, the corresponding blending noise can be predicted and subtracted from the raw pseudo-deblended seismic data, and thus an updated input dataset with considerably weakened blending noise is obtained. The main goal of the network training in subsequent steps shifts to attenuating weak blending noise and extracting more detailed signal features leaking from previous steps. Hence, the optimized model in the previous deblending step is used as an initialization for the fine-tuning of the current step through transfer learning, which is more efficient than a random initialization. In addition, the loss function is also slightly modified using a structural similarity (SSIM) loss in the subsequent training steps to better attenuate the remaining blending noise. The introduced multistep deblending algorithm converges quickly, and for efficiency and storage considerations, we set the number of steps to three in this work. Finally, with the optimized parameters of different individual networks, the multistep deblending can be sequentially implemented on CRGs of the raw pseudo-deblended data. Due to the coherence similarity between CRGs and CSGs of pseudo-deblended data, the optimized network using the CSGs can generate satisfactory deblended results of the CRGs for subsequent seismic processing. Numerical examples of synthetic data indicate the validity and flexibility of the proposed self-supervised multistep algorithm, and applications to field artificially blended data demonstrate its adaptability in improving deblending accuracy without requiring clean labels.

2 Theory

2.1 Review on Blended Acquisition

For efficient seismic data acquisition, the blended acquisition has drawn much interest and is still open compared to the conventional acquisition algorithm. For a specific receiver $r_i, i = 1, 2, \dots, Nr$, it records signals continuously from multiple sources fired almost simultaneously, and the continuously recorded blended data d_{bl} can be illustrated mathematically as,

$$d_{bl}(t, r_i) = \sum_{j \in S} d(t + \tau_j, s_j, r_i), i = 1, 2, \dots, Nr, \quad (1)$$

where Nr denotes the number of receivers. The dithering time τ_j in blended sources $s_j, j \in S$ can ensure that the blending noise is randomized in certain domains, like CRD and CMD. Figure 1 depicts a comparison cartoon of the conventional acquisition and the blended acquisition with a blending fold of two. When two blended sources are fired almost simultaneously, the receivers continuously record the signal from the two blended sources.

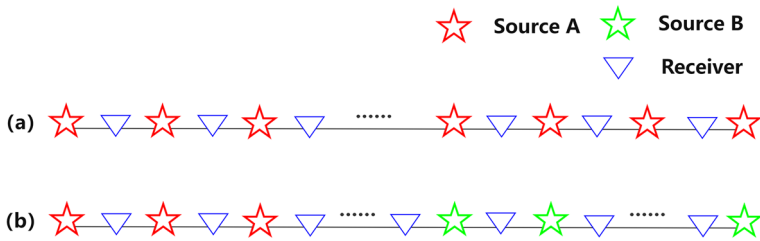


Fig. 1 **a** The conventional seismic acquisition; **b** Blended acquisition to improve the efficiency by nearly twice with a blending fold of two

Compared to conventional acquisition with a large firing time interval to avoid signal interference, this blended acquisition offers an approximately twofold increase in acquisition efficiency. The blending fold, which is proportional to the ratio between the conventional acquisition period and the blended acquisition period, reflects the improvement in efficiency achieved through blended acquisition.

The blending formula in Eq. (1) can be reformulated as a compact formula using the matrix–vector symbol,

$$\mathbf{d}_{bl} = \mathbf{\Gamma} \mathbf{d} \tag{2}$$

where \mathbf{d}_{bl} represents continuously recorded data, $\mathbf{\Gamma}$ represents the blending operator determined with the dithering time of blended sources, and \mathbf{d} represents the conventional unblended record in the CRD. Then, the corresponding pseudo-deblended data \mathbf{d}_{pdb} is obtained using the adjoint blending operator $\mathbf{\Gamma}^H$ (Lin and Sacchi 2020),

$$\mathbf{d}_{pdb} = \mathbf{\Gamma}^H \mathbf{d}_{bl} = \mathbf{d} + (\mathbf{\Gamma}^H \mathbf{\Gamma} - \mathbf{I}) \mathbf{d}. \tag{3}$$

Pseudo-deblended data \mathbf{d}_{pdb} is the sum of the unblended record \mathbf{d} and the blending noise $(\mathbf{\Gamma}^H \mathbf{\Gamma} - \mathbf{I}) \mathbf{d}$, and the symbol \mathbf{I} represents the identity operator. To visually illustrate the blending process and the effects of the blending noise, a synthetic record in the CRD is used for illustration. Figure 2a depicts an unblended record. A blending fold of two is used for blended acquisition (Fig. 1b) with the dithering time located in $[-0.25 \text{ s}, 0.25 \text{ s}]$, and Fig. 2b illustrates the corresponding pseudo-deblended data, where the randomized blending noise negatively contaminates the coherent signal. Using a coherence-pass filtering or inversion algorithm, the randomized blending noise can be attenuated, and then deblended results are obtained.

2.2 Supervised Multistep Deblending with Quantitative Blending Noise Evaluation

Deep learning-based algorithms accurately characterize seismic data by extracting high-level, nonlinear features from training data. The deblending accuracy and efficiency are high when the designed supervised network is trained or fine-tuned adequately. The commonly used loss function for training the network, denoted as Eq. (4), aims to optimize the network parameters θ ,

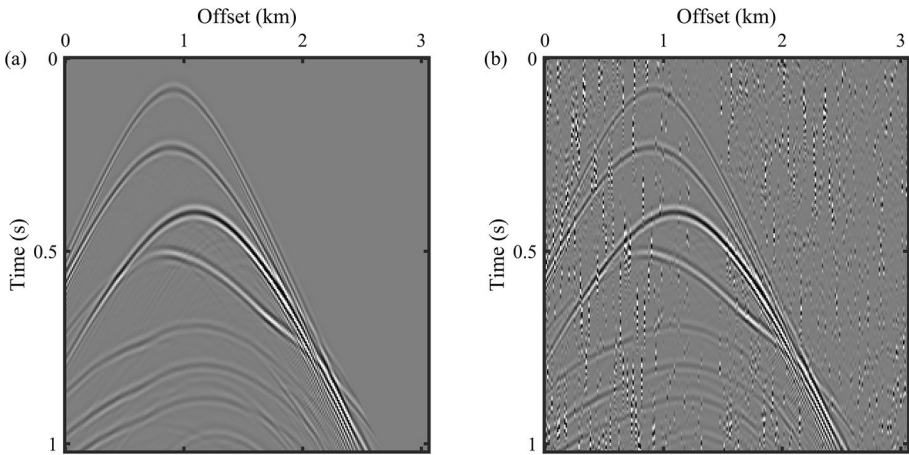


Fig. 2 a Unblended seismic record; b The corresponding pseudo-deblended record in the CRD

$$L_1(\theta) = \|f(\mathbf{d}_{\text{pdb}}; \theta) - \mathbf{d}\|_2^2 + \lambda \|\mathbf{W}\|_2^2, \tag{4}$$

where f represents the designed U-net architecture (Fig. 3) used in the paper. The symbol \mathbf{W} represents the smoothness constrained convolution kernel, and λ is a balance factor between the smoothness of the convolution kernel and the L_2 norm measured data misfit. Using sets of labeled training data, the optimized parameters θ_1^* can be determined when the training converges, during which the Adam optimizer is used. To quantitatively evaluate the deblending performance, the signal-to-noise ratio (SNR) is introduced,

$$\text{SNR}(\text{dB}) = 10 \log_{10} \frac{\|\mathbf{d}\|^2}{\|\mathbf{d} - \mathbf{d}_{\text{dbl}}\|^2}, \tag{5}$$

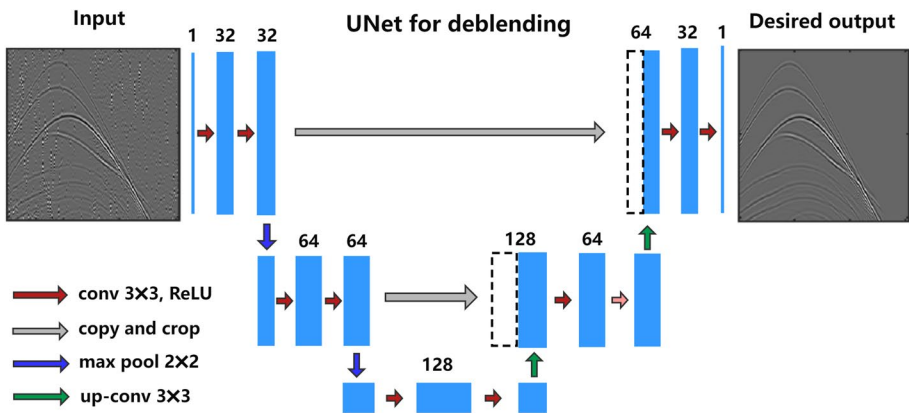


Fig. 3 The designed U-net architecture for supervised deblending

where \mathbf{d} represents the unblended record and \mathbf{d}_{dbl} represents the deblended result. With the optimized parameters θ_1^* , the deblended data $\mathbf{d}_{\text{dbl}}^1$ is obtained by feeding pseudo-deblended data $\mathbf{d}_{\text{pdb}} = \Gamma^H \mathbf{d}_{\text{bl}}$ to the optimized network f ,

$$\mathbf{d}_{\text{dbl}}^1 = f(\mathbf{d}_{\text{pdb}}, \theta_1^*). \tag{6}$$

Since the blending noise can be precisely predicted with the coherent signal and shot time, an iterative blending noise estimation–subtraction strategy (Mahdad et al. 2011) is used to improve the deblending accuracy. The simulated blending noise $(\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^1$ can be obtained using the deblended result $\mathbf{d}_{\text{dbl}}^1$ in the first deblending step, and the input of the next deblending step is updated to $\mathbf{d}_{\text{pdb}} - (\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^1$ after the blending noise estimation–subtraction procedure. The updated input contains only weak blending noise contamination. The well-trained network with the optimized parameters θ_1^* , however, cannot characterize the weak blending noise accurately because the training procedure does not recognize the weak blending noise (Wang et al. 2021a). Further deblending accuracy improvement requires how to quantitatively evaluate the remaining blending noise in a recursive strategy, and a multistep deblending algorithm is adopted. Since the optimized parameters θ_1^* attenuate dominant blending noise, the fine-tuning step is used for the remaining weak blending noise evaluation via transfer learning with θ_1^* as an initialization. The input is updated to $\mathbf{d}_{\text{pdb}} - (\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^1$, and the label \mathbf{d} remains the same. After the fine-tuning convergence, the optimized parameters θ_2^* are obtained to achieve deblended results with the remaining weak blending noise attenuation, i.e., $\mathbf{d}_{\text{dbl}}^2 = f(\mathbf{d}_{\text{pdb}} - (\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^1; \theta_2^*)$ via the second deblending step. For subsequent deblending steps, we quantitatively evaluate the remaining weak blending noise in a multistep manner with the training pairs $(\mathbf{d}_{\text{pdb}} - (\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^j, \mathbf{d})$, $j = 2, 3, \dots, J$, and then the multistep deblending is implemented in a recursive manner with each step employing an individual network with optimized parameters.

The multistep deblending algorithm here includes a slight modification to the loss function to improve the deblending performance according to different levels of the remaining blending noise in each deblending step, which exploits the flexibility of the multistep algorithm. Since the dominant signal has been recovered in the first deblending step, it is more important to retain the detailed structure of the coherent signal when attenuating the remaining weak blending noise. As a measurement of the similarity between two images, the SSIM loss function is incorporated into DL-based seismic processing for better performance (Li et al. 2021; Liu et al. 2022; Yu and Wu 2022). Thus, we introduce the SSIM loss (Eq. (7)) into the raw loss function (Eq. (4)) to recover the details of deblended results when suppressing the remaining blending noise,

$$L_{\text{SSIM}} = 1 - \text{SSIM}(f(\mathbf{d}_{\text{pdb}} - (\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^j; \theta), \mathbf{d}), \quad j = 1, 2, \dots, J. \tag{7}$$

The final loss function in fine-tuning steps can be formulated as,

$$L_{j+1}(\theta) = \mu \left\| f(\mathbf{d}_{\text{pdb}} - (\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^j; \theta) - \mathbf{d} \right\|_2^2 + (1 - \mu)L_{\text{SSIM}} + \lambda \|\mathbf{W}\|_2^2, \quad j = 1, 2, \dots, J, \tag{8}$$

where μ is a weighting parameter ranging in $[0, 1]$, balancing the data misfit and the SSIM loss. The input $\mathbf{d}_{\text{pdb}} - (\Gamma^H \Gamma - \mathbf{I})\mathbf{d}_{\text{dbl}}^j$, $j = 1, 2, \dots, J$ and the desired output \mathbf{d} are used to fine-tune the previously optimized parameters θ_j^* via transfer learning, and the fine-tuned

parameters θ_{j+1}^* are obtained to attenuate the remaining blending noise in a multistep manner. Using the above-mentioned multiple steps, we can obtain different optimized parameters θ_j^* , $j = 1, 2, \dots, J$ of individual networks. The blending noise can be attenuated step by step to achieve accurate debledned results for the test pseudo-debledned data \tilde{d}_{pdb} because of its similar features to the training data,

$$\begin{aligned} \tilde{d}_{dbl}^1 &= f_1(\tilde{d}_{pdb}; \theta_1^*) \\ \tilde{d}_{dbl}^2 &= f_2(\tilde{d}_{pdb} - (\Gamma^H \Gamma - \mathbf{I})\tilde{d}_{dbl}^1; \theta_2^*) \\ &\dots \\ \tilde{d}_{dbl}^{J+1} &= f_{J+1}(\tilde{d}_{pdb} - (\Gamma^H \Gamma - \mathbf{I})\tilde{d}_{dbl}^J; \theta_{J+1}^*) \end{aligned} \tag{9}$$

Figure 4 summarizes the workflow of the multistep (three-step) debledning algorithm with input details during different individual network training and subsequent debledning. However, supervised debledning requires clean labels, which are rarely available in field cases. In field cases, the debledned results from conventional debledning algorithms are usually regarded as labels (Wang et al. 2023), and the inaccuracy of debledned results negatively impacts the effectiveness of DL-based debledning methods. Therefore, constructing the training labels accurately and flexibly becomes essential.

2.3 Self-Supervised Multistep Debledning

In contrast to the blending noise, which is coherent in the CSD but is spatially randomized in the CRD, the signal of pseudo-debledned data is coherent in both CRD and CSD. The coherent features, such as dip, curvature, and amplitude, are almost consistent in these two domains based on Green’s reciprocity theorem of the wave equation if all source wavelets are consistent (Han et al. 2022; Wang and Hu 2022). A pseudo-debledned CRG with the coherent signal and randomized blending noise is depicted in Fig. 5a. Figure 5b demonstrates a CSG of the pseudo-debledned data with coherent blending noise and signal. The coherent characteristics of CRG and CSG of pseudo-debledned data are similar, which qualitatively illustrates Green’s reciprocity theorem of the wave equation. We present a

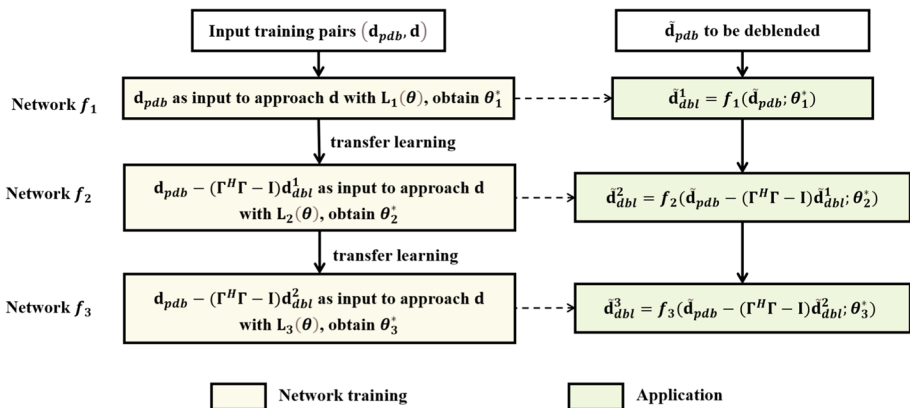


Fig. 4 The summarized workflow of the multistep (three-step) debledning algorithm

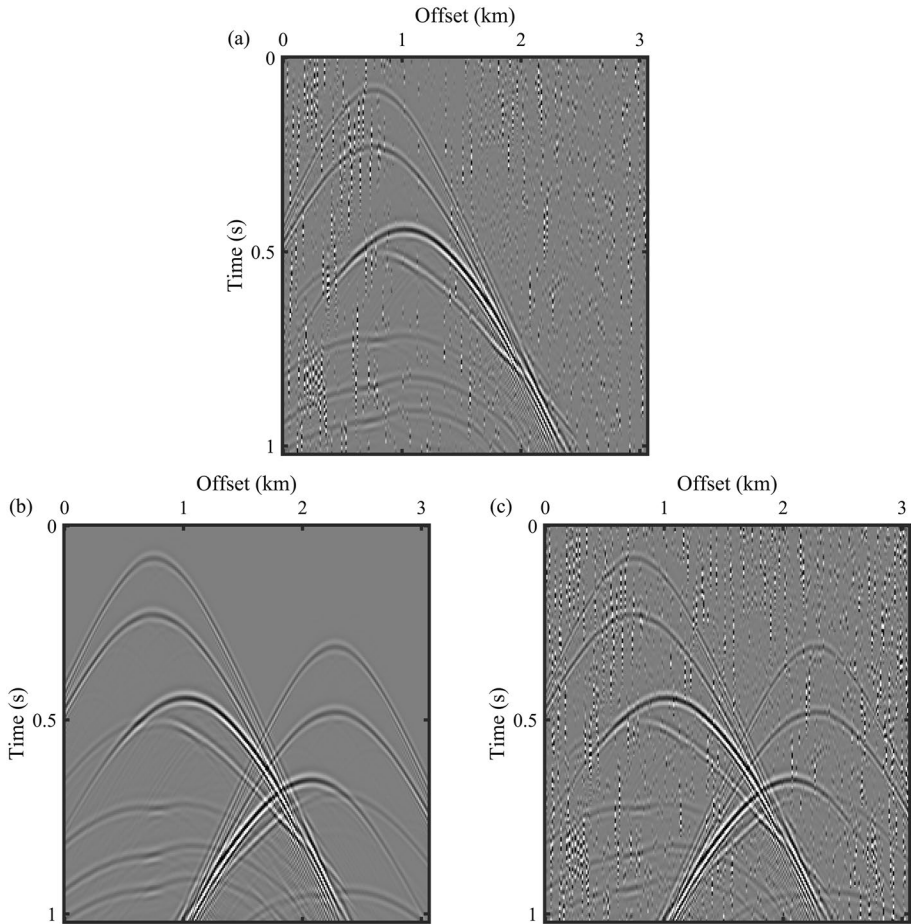


Fig. 5 **a** The CRG of pseudo-deblended records with a blending fold of two; **b** The corresponding CSG of pseudo-deblended records; **c** The constructed data with the additional blending noise contamination based on **(b)**

self-supervised multistep deblending algorithm based on the similarity between the CSGs and CRGs of the raw pseudo-deblended data.

We use the CSGs of the raw pseudo-deblended data to construct the adaptive training datasets. The CSGs \mathbf{d}_{csg} , with coherent blending noise, are regarded as the labels (depicted in Fig. 5b). A blending operator and its adjoint operator are used to construct the blending noise-contaminated data $\mathbf{d}_{\text{csg_pdb}} = \mathbf{\Gamma}^H \mathbf{\Gamma} \mathbf{d}_{\text{csg}}$, as depicted in Fig. 5c with details. Thus, we can obtain the adaptive training data pairs $(\mathbf{d}_{\text{csg_pdb}}, \mathbf{d}_{\text{csg}})$ for a supervised algorithm, generating a self-supervised algorithm as the training data is constructed from the raw blended data itself. A self-supervised multistep algorithm can capture the nonlinear relationship between the input data with varying levels of the remaining blending noise and the labels using different individual networks, as illustrated in the previous subsection. Using the proposed self-supervised multistep deblending algorithm, we can obtain the optimized parameters $\theta_j^*, j = 1, 2, \dots, J$ for different networks after training convergence using adaptively constructed training data in a stepwise

manner. The parameters $\theta_j^*, j = 1, 2, \dots, J$ are used to deblend the raw pseudo-deblended CRGs (depicted in Fig. 5a) in a sequential and multistep manner. The deblending accuracy can be guaranteed because of the similar coherence of the training CSGs and the test CRGs of pseudo-deblended data.

To fully illustrate the proposed algorithm, a detailed workflow is summarized in Fig. 6. First, we extract the CSGs of pseudo-deblended data as training labels. Next, we obtain the training input with additional blending noise contamination by numerically blending and pseudo-deblending, generating adaptive training pairs $(\mathbf{d}_{csg_pdb}, \mathbf{d}_{csg})$ for the multistep deblending algorithm. During the multistep training procedures, the training input is updated via blending noise estimation–subtraction, and the label of each network is the CSGs of the raw pseudo-deblended data. The introduced multistep deblending algorithm converges very fast, and we employ three steps to attenuate the dominant, weak and even weaker blending noise in a quantitative manner for storage and efficiency considerations. During the fine-tuning procedure, the optimized model from the previous step initializes the current individual network. Finally, the optimized parameters $\theta_j^*, j = 1, 2, \dots, J$ of individual networks are sequentially employed to deblend CRGs of the raw pseudo-deblended data in multiple steps. The deblended results can be obtained accurately without unblended label requirements. The effectiveness of the proposed deblending algorithm is shown in the following section using synthetic and field blended data examples.

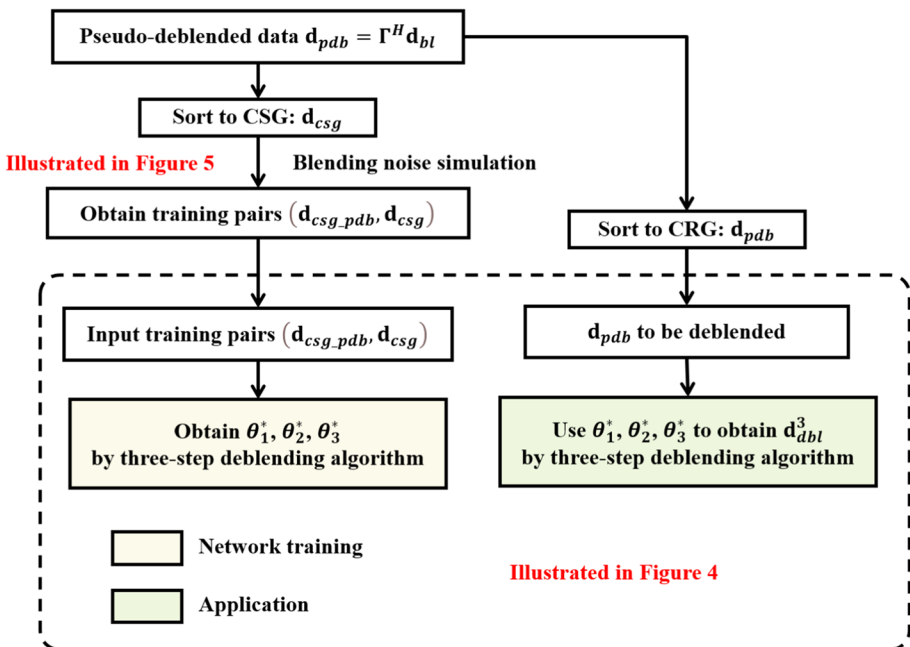


Fig. 6 The workflow of the self-supervised multistep (three-step) deblending algorithm

3 Numerical Examples

To illustrate the superiority of the proposed self-supervised multistep deblending algorithm, numerically blended synthetic and field data examples are presented, along with comparisons to the MBNS by Wang et al. (2022a).

3.1 Synthetic Data Example

The synthetic data is simulated from a layered model with a high-velocity salt body using finite-difference forward modeling, and it includes 256 shots and 256 receivers located uniformly. The spatial interval is set to 12 m, and the sampling time interval is set to 4 ms during forward modeling. As a deblending reference, Fig. 7a depicts a specific unblended CRG record. With a blending fold of four, we obtain artificially blended data, and the CRG of the corresponding pseudo-deblended data is extracted, as depicted in Fig. 7b. Randomized blending noise (Fig. 7c) seriously contaminates the coherent signal.

A CSG record of pseudo-deblended data is depicted in Fig. 7d, and the coherent signal is contaminated by coherent interference. As previously mentioned, we can use CSGs of pseudo-deblended data \mathbf{d}_{csg} as labels to simulate the blending noise-contaminated data via $\mathbf{d}_{\text{csg_pdb}} = \mathbf{\Gamma}^H \mathbf{\Gamma} \mathbf{d}_{\text{csg}}$, as illustrated in Fig. 7e. Then, we can obtain the adaptive training sets $(\mathbf{d}_{\text{csg_pdb}}, \mathbf{d}_{\text{csg}})$. Figure 7f shows the constructed blending noise, and it is consistent with the raw blending noise (Fig. 7c). Thus, the optimized parameters trained using the constructed adaptive training sets $(\mathbf{d}_{\text{csg_pdb}}, \mathbf{d}_{\text{csg}})$ from the CSGs

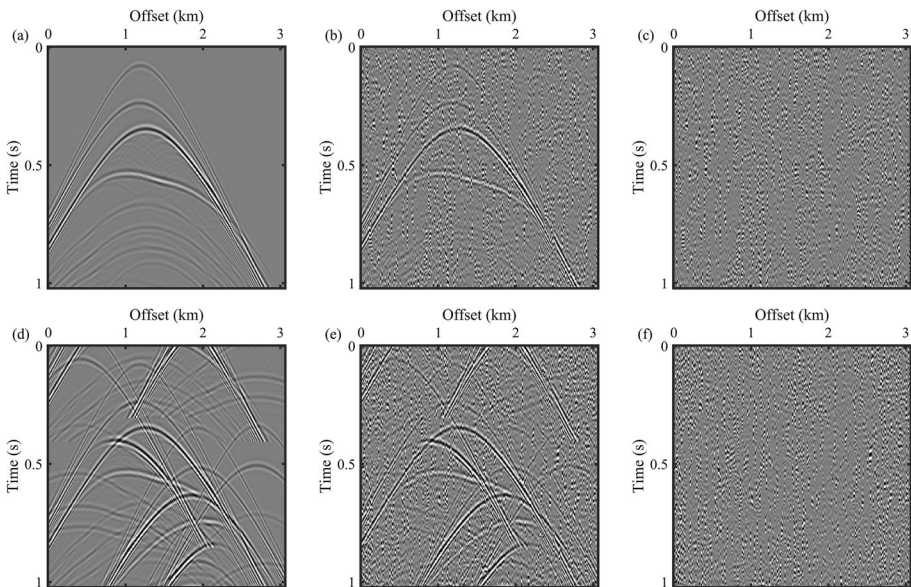


Fig. 7 The comparison of the test data and the constructed training data. **a** The unblended synthetic record; **b** The corresponding pseudo-deblended test CRG data; **c** The difference between **(b)** and **(a)**, i.e., the raw blending noise; **d** The corresponding CSG of pseudo-deblended data, i.e., the constructed training label; **e** The constructed data with the additional blending noise contamination based on **(d)**, i.e., the constructed initial training input; **f** The difference between **(e)** and **(d)**, i.e., the constructed blending noise

can accurately attenuate the CRG blending noise of the raw pseudo-deblended data. With 224 training pairs and 32 validation pairs from the adaptively constructed training data, we can get the optimized parameters of three individual networks using the multistep deblending algorithm.

3.1.1 Training on the CSGs of Pseudo-deblended Data

The number of training epochs for the first, second, and third individual networks in the proposed self-supervised multistep algorithm is set to 400, 200, and 150, respectively, and the total number of training epochs is 750. The first deblending step tries to attenuate the dominant blending noise and obtain a preliminary deblended result for subsequent estimation–subtraction of the blending noise. Figure 8a, b depicts the input of the first deblending step and the corresponding blending noise. After the blending noise estimation–subtraction process using the deblended result of the first step, we update the input of the second step with $\mathbf{d}_{\text{csg_pdb}} - (\mathbf{\Gamma}^H \mathbf{\Gamma} - \mathbf{I}) \mathbf{d}_{\text{csg_dbl}}^1$. Figure 8c illustrates the training input of the second deblending step, containing the remaining weak blending noise to be attenuated and signal leakage from the first step (Fig. 8d). In the third deblending step, the even weaker signal leakage and the remaining blending noise from the second deblending step can be processed. Figure 8e depicts the training input of the third network for deblending, and Fig. 8f depicts the corresponding residual with the label, where the blending noise and signal leakage are even weaker. From the sight of the remaining blending noise in the sequential deblending inputs (Fig. 8b, d and f), we can see its decreasing trend. The proposed algorithm can evaluate the decreasing blending noise quantitatively using different individual networks in a stepwise manner. In the self-supervised multistep deblending algorithm, the optimized parameters of one network can be treated as an initialization of the subsequent network via transfer learning, with the same label for different individual networks.

Figure 9a displays the recovered SNR curves of the training and validation data using the proposed self-supervised multistep deblending algorithm during sequential training. The SNR curve of the training data (the blue curve) exhibits a pattern similar to that of the validation data (the red curve) but is slightly higher. The proposed deblending algorithm consists of individual training of three sequential networks, and the averaged SNR increases step by step. To illustrate the superiority of the proposed algorithm, the self-supervised MBNS is used for detailed comparisons, using predefined scalars α for the training dataset construction. To simulate the decreasing blending noise level during the used three deblending iterations, three scalars are set to $\alpha = 1.0, 0.25, 0.0001$, respectively, by trial and error to fulfill and diverge the training set. Then, the training data pairs become $(\mathbf{d}_{\text{csg}} + \alpha(\mathbf{\Gamma}^H \mathbf{\Gamma} - \mathbf{I}) \mathbf{d}_{\text{csg}}, \mathbf{d}_{\text{csg}})$ according to three predefined blending noise levels; thus, the amount of training data triples. To keep the training burden similar, the training epoch number for the MBNS is set to 250. Figure 9b illustrates the recovered SNRs as the increasing training epochs, where the blue curve represents the recovered SNR of the training data and the solid red curve represents that of the validation data. As the attenuation of the raw blending noise with a scaler $\alpha = 1.0$ is the main task for the deblending issue, we list the recovered SNR of the validation data, as marked by the dashed red curve.

All validation datasets are processed using the above-mentioned two strategies to validate the proposed self-supervised multistep algorithm. Figure 10a depicts the recovered SNRs, with the red curve representing that of the multistep algorithm and the blue curve indicating that of the MBNS (three-level) with three iterations. The proposed multistep (three-step) algorithm recovers a higher average SNR (20.5 dB) than the MBNS (18.9 dB).

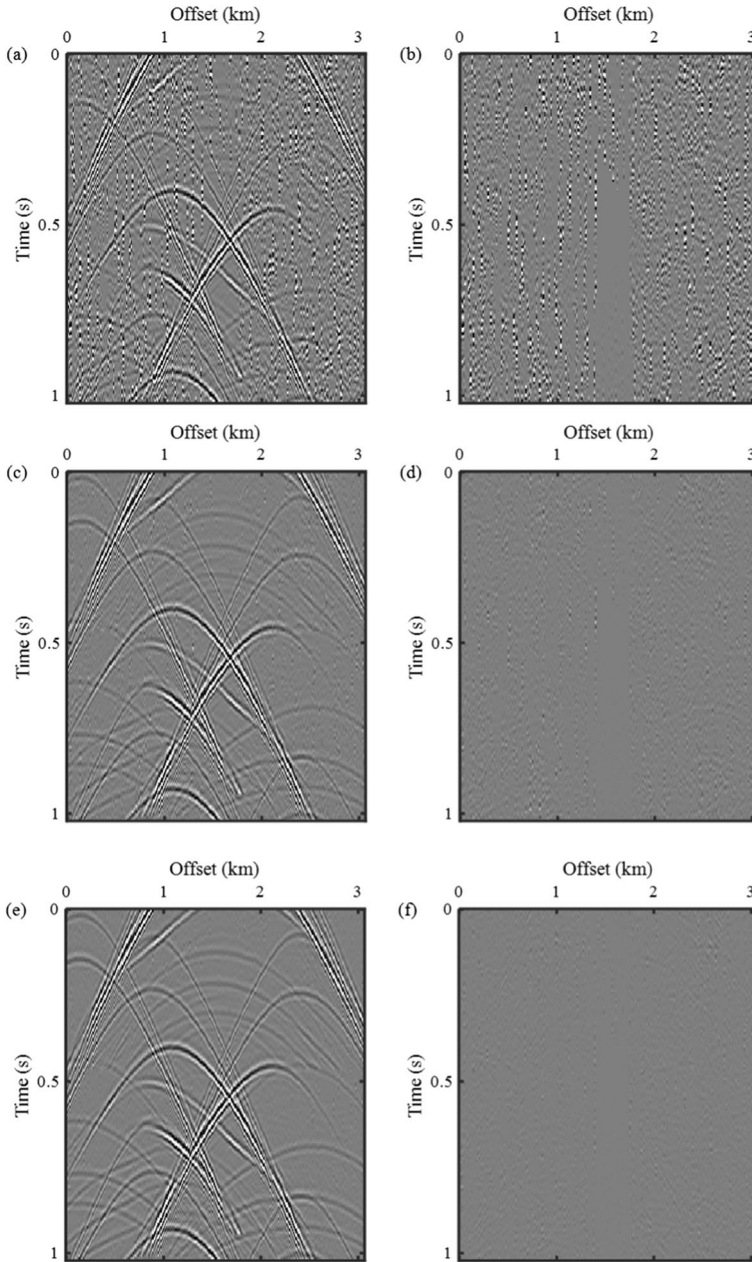


Fig. 8 a, c, e The input of the sequential first, second, and third network for debloating; b, d, f The residual between (a, c, e), and the training label, i.e., the blending noise and signal leakage

Another evaluation index, i.e., SSIM, is evaluated as an average of 0.92 and 0.85, respectively, as shown in Fig. 10b, for the debloating performance assessment. From the performance analysis of the validation data debloating, the multistep (three-step) debloating

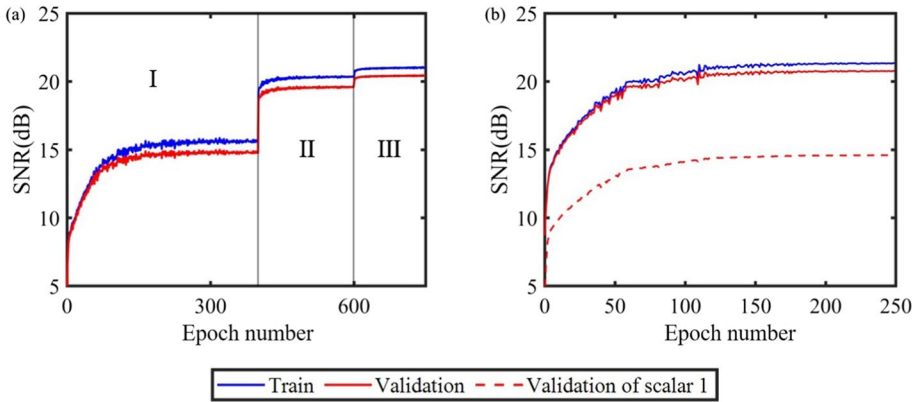


Fig. 9 The recovered SNR convergence curve during the network training of **a** the proposed multistep (three-step) algorithm and **b** the MBNS (three-level)

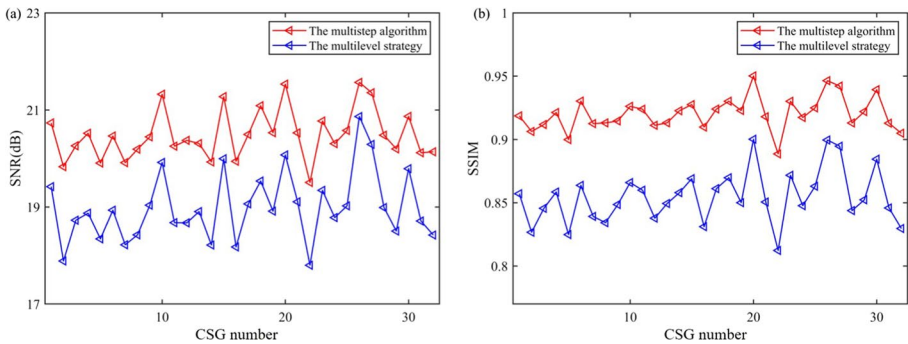


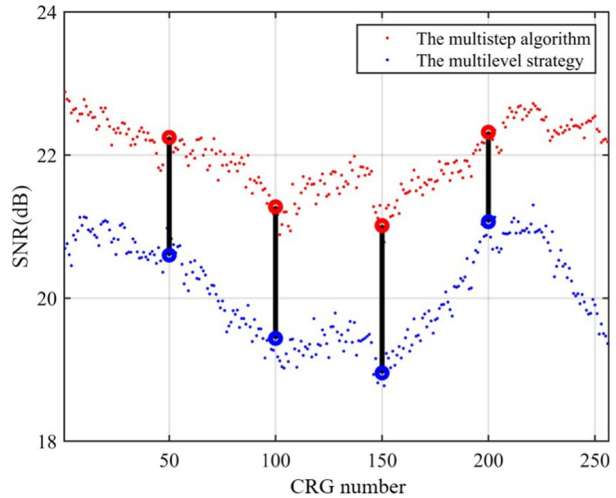
Fig. 10 The deblending performance evaluation using **a** the recovered SNR and **b** the recovered SSIM of the validation data using the proposed self-supervised multistep (three-step) algorithm and the self-supervised MBNS (three-level) with three iterations

algorithm is superior to the MBNS (three-level) with three iterations, which positively supports our proposed multistep algorithm.

3.1.2 Deblending the CRGs of Pseudo-deblended Data

After the training convergence, we use the optimized networks for the CRG deblending of the raw pseudo-deblended data. All 256 test CRGs are deblended via the above-mentioned two methods for deblending performance comparison, and the recovered SNRs are depicted by the red and blue dots in Fig. 11. The performance of the 50th, 100th, 150th, and 200th CRGs is specifically marked by circles to visualize that the multistep deblending algorithm significantly improves the recovered SNR compared to the MBNS (three-level) with three iterations. The comparisons indicate that the deblended results obtained using the multistep (three-step) algorithm outperform those of the MBNS (three-level) with three iterations, and the average recovered SNRs are 21.98 dB and 20.09 dB, respectively. Figure 11 also

Fig. 11 The recovered SNRs of all test CRGs of the raw pseudo-deblended data by the multistep (three-step) algorithm (red dots) and the MBNS (three-level) with three iterations (blue dots). The performance of the 50th, 100th, 150th, and 200th CRGs is marked by circles



validates the self-supervised multistep deblending algorithm, which uses CSGs constructed training datasets to achieve the optimized networks for deblending CRGs.

To further illustrate the superior deblending performance of the multistep-assisted self-supervised algorithm, we extract a CRG record from the raw pseudo-deblended data and enlarge the complex part, specifically focusing on diffraction details. Figure 12a shows the 100th CRG of the unblended record, i.e., the reference during the deblending performance evaluation. Figure 12b depicts the enlarged complex part that is marked by the red box in Fig. 12a and c shows the corresponding pseudo-deblended CRG data. After the three sequential deblending steps of the proposed self-supervised multistep algorithm, the final deblended result is obtained with a recovered SNR of 21.28 dB in the global part and a recovered SNR of 17.65 dB in the enlarged complex part. Figure 12d–f shows the deblended result, the enlarged complex part of the deblended result, and the residual using the multistep algorithm. Figure 12g shows the deblended result obtained using the MBNS (three-level) with three iterations, and the recovered SNR is approximately 1.5 dB lower. Furthermore, the recovered SNR in the enlarged complex part (Fig. 12h) is approximately 2.7 dB lower compared with that of the multistep (three-step) deblending algorithm. The residual comparison in Fig. 12f, i also validates the deblending performance of the self-supervised multistep algorithm in complex local areas, which is superior to the MBNS in recovering weak diffractions with approximately 3 dB in terms of SNR improvement.

By analyzing the deblended results in the complex local part and their corresponding residuals, we find that the multistep algorithm outperforms the MBNS in terms of weak signal preservation and blending noise attenuation. The multistep deblending algorithm quantifies the remaining blending noise level after the blending noise estimation–subtraction during each deblending step instead of using a fixed scalar to qualitatively describe the blending noise level in the MBNS. Besides, the multistep algorithm effectively extracts the signal leakage from the previous deblending step, allowing for the recovery of more detailed signals, especially weak diffractions. In addition, transfer learning is employed in the training of adjacent individual networks to enhance training efficiency. Detailed comparisons of the synthetic examples show the superiority of the proposed algorithm, which constructs adaptive training data based

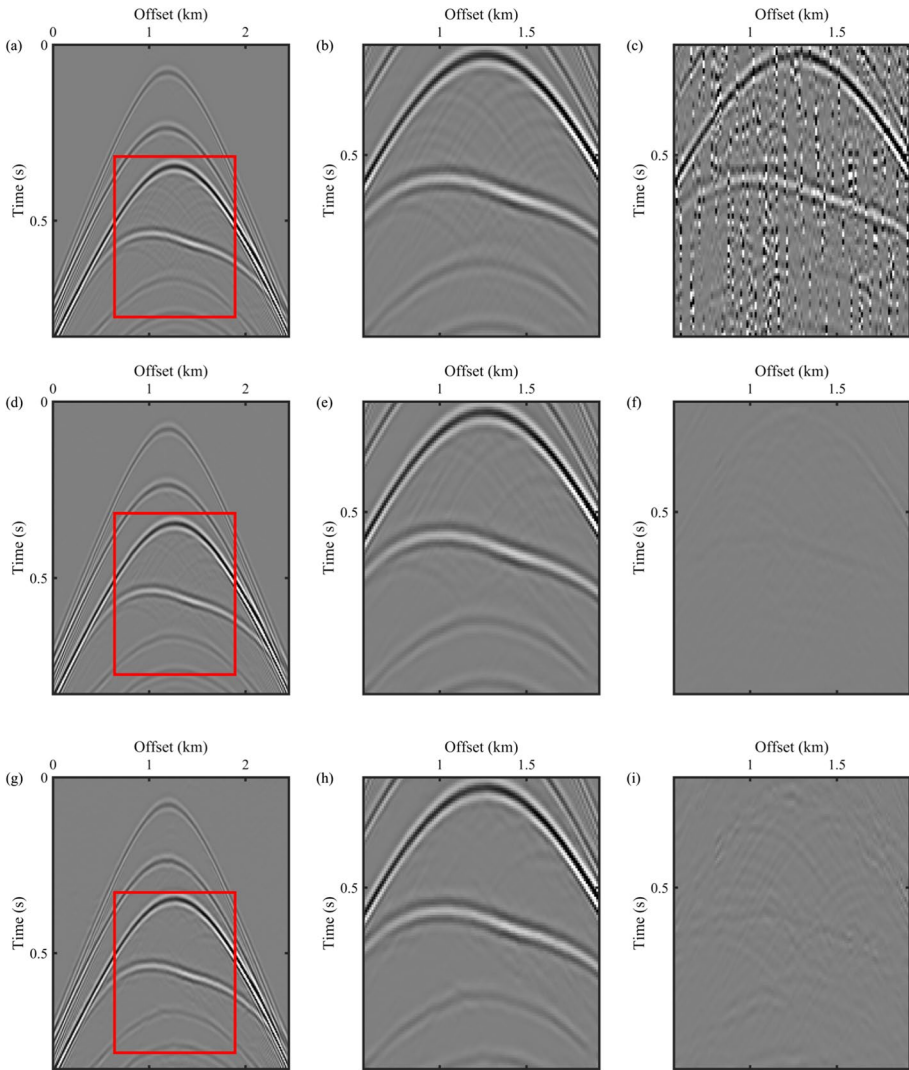


Fig. 12 Deblending performance comparison of the 100th CRG. **a** The reference unblended record; **b** The enlarged part of the red box in **(a)**; **c** The corresponding pseudo-deblended record of **(b)**; **d** Deblended result of the multistep (three-step) deblending algorithm; **e** The enlarged part of the red box in **(d)**; **f** The residual between **(b)** and **(e)**; **g** Deblended result of the MBNS (three-level) with three iterations; **h** The enlarged part of the red box in **(g)**; **i** The residual between **(b)** and **(h)**

on raw pseudo-deblended CSGs and deblends the raw pseudo-deblended CRGs with the optimized individual networks in a stepwise manner. Next, we use the adaptively constructed CSG training data of field data to fine-tune the first optimized network of synthetic data based on transfer learning, and the optimized multiple networks can be obtained to deblend the CRGs of raw pseudo-deblended field data using the multistep deblending algorithm.

3.2 Field Data Example

We have processed a 2D line survey with 128 shots, 128 geophones per shot, and 512 samples in each trace. The preprocessed field data is from the Gulf of Suez, with a sampling time interval of 4 ms and a spatial interval of 12.5 m. The unblended field data is numerically blended with a blending fold of three to simulate field blended data acquisition. A specific CRG of unblended field data is shown in Fig. 13a, and the corresponding CRG of the pseudo-deblended data is shown in Fig. 13b. The coherent signal (Fig. 13a) is negatively contaminated by the randomized blending noise (Fig. 13c). To construct adaptive training sets for self-supervised algorithms in field data applications, we extract the CSGs from pseudo-deblended field data, and one of them is shown in Fig. 13d. Then, we artificially construct blending noise-contaminated seismic data as shown in Fig. 13e, which is regarded as the initial training input with the CSGs in Fig. 13d as the desired output. The constructed blending noise in Fig. 13f is similar to the raw blending noise in Fig. 13c. The coherence consistency between Fig. 13b and d guarantees the accuracy of the final deblending results for the CRGs of the raw pseudo-deblended data using the optimized network with the CSG-assisted adaptive training sets. The multistep deblending algorithm uses 128 pairs of adaptive training data (Fig. 13d, e) as the desired output and initial training input.

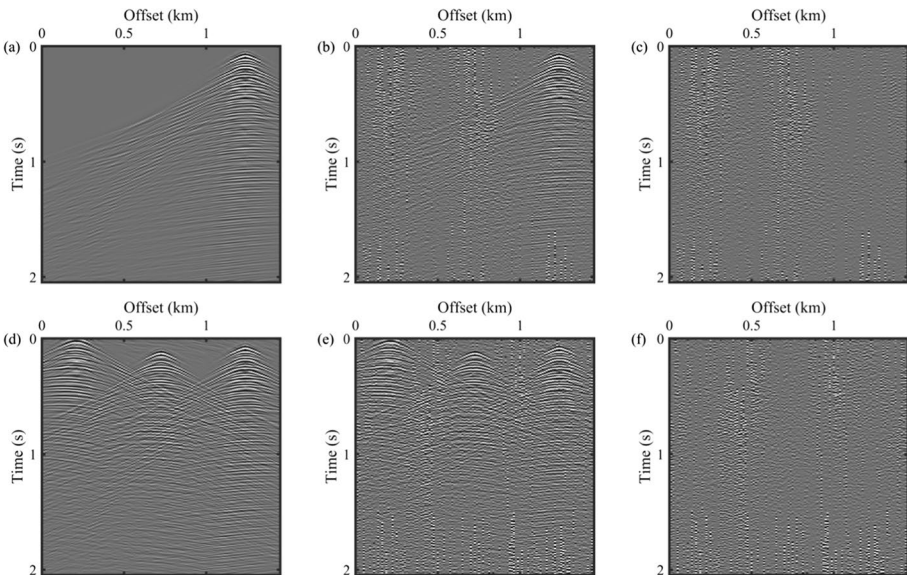


Fig. 13 The comparison of the test data and the constructed training data. **a** The unblended field record; **b** The corresponding pseudo-deblended test CRG data; **c** The difference between **(b)** and **(a)**, i.e., the raw blending noise; **d** The corresponding CSG of pseudo-deblended data, i.e., the constructed training label; **e** The constructed data with the additional blending noise contamination based on **(d)**, i.e., the constructed initial training input; **f** The difference between **(e)** and **(d)**, i.e., the constructed blending noise

3.2.1 Training on the CSGs of Pseudo-deblended Field Data

For deblending field data, we fine-tune the first optimized network θ_1^* of synthetic data using 112 pairs of field training data via transfer learning. The remaining 16 field data pairs are used as the validation data. The training epochs of three networks are set to 200, 100, and 60, respectively, for the proposed self-supervised multistep (three-step) deblending algorithm. The input of the first deblending network is seismic data with strong blending noise contamination (Fig. 14a). Figure 14b, c shows the input of the subsequent deblending networks as the deblending step increases, which contains gradually weaker blending noise. Figure 14d, e, and f shows the residual between the deblending input and the training label, which consists of the remaining blending noise and signal leakage from the previous step.

To illustrate the superiority of the self-supervised multistep algorithm on field data applications, the MBNS with a similar computation burden is implemented for detailed comparisons. Three scalars (1, 0.3, 0.0001) are used for the training data construction, and the training epoch number for the MBNS (three-level) is set to 120. Figure 15 illustrates the recovered average SNR with the increasing training epoch using the proposed self-supervised multistep (three-step) deblending algorithm (Fig. 15a) and the MBNS (Fig. 15b) during network training. The blue curve represents the recovered SNR of the training data, and the solid red curve represents that of the validation data. The dashed red curve represents the recovered SNR of the validation data with the raw blending noise level (i.e., the scalar $\alpha = 1$) using the MBNS.

3.2.2 Deblending the CRGs of Pseudo-deblended Field Data

When the training converges, we can obtain the optimized network parameters to deblend all CRGs of the raw pseudo-deblended field data in a stepwise manner. For a better

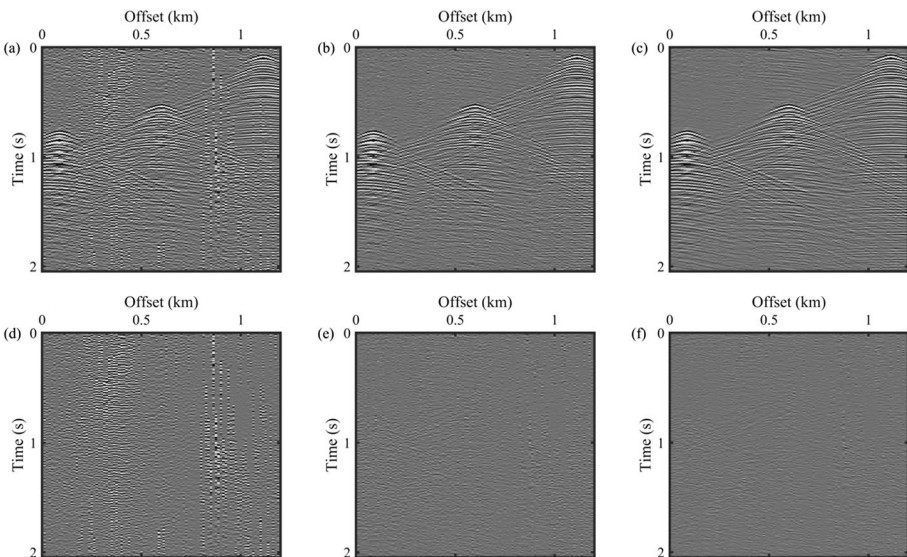


Fig. 14 **a, b, c** The input of the sequential first, second, and third network for deblending; **d, e, f** The residual between **(a, b, c)** and the training label, i.e., the blending noise and signal leakage

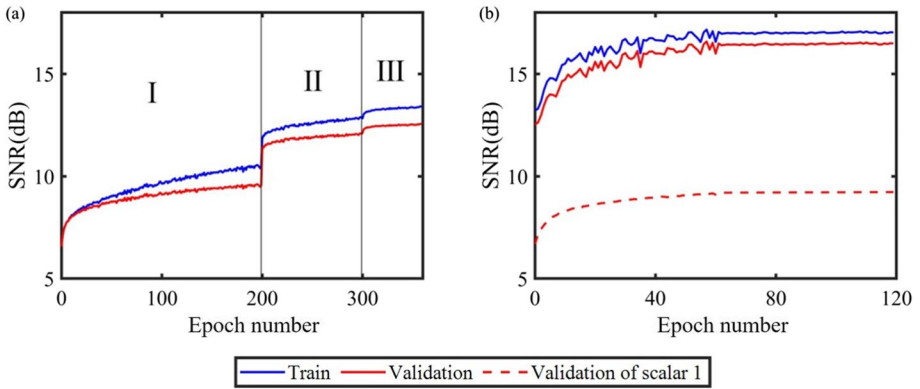


Fig. 15 The recovered SNR convergence curve during the network training of (a) the proposed self-supervised multistep (three-step) algorithm and (b) the MBNS (three-level)

comparison, deblending is also performed using the MBNS (three-level) with three iterations. Figure 16 shows the recovered SNRs of each test CRG when the deblended data is compared with the referenced unblended data. The result of the proposed algorithm is depicted by the red dots, and that of the MBNS with three iterations is depicted by the blue dots. The average recovered SNR is 12.56 dB for deblended results using the proposed multistep algorithm. The MBNS achieves an average recovered SNR of 10.88 dB after three iterations, which is approximately 1.7 dB lower than the multistep algorithm. Figure 16 shows that the multistep algorithm has a significant improvement in deblending performance compared to the MBNS during CRG deblending. The reason is that the multistep deblending algorithm quantifies the remaining blending noise levels using different

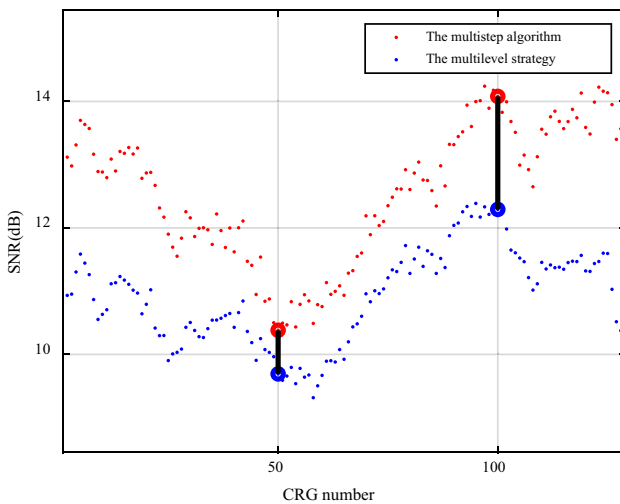


Fig. 16 The recovered SNRs of all test CRGs of the raw pseudo-deblended data by the multistep (three-step) algorithm (red dots) and the MBNS (three-level) with three iterations (blue dots). The performance of the 50th and 100th CRGs is marked by the circles

individual networks. In contrast, the MBNS employs scalars to qualitatively imitate the blending noise levels, which inevitably biases the true blending noise levels during iterative deblending procedures. The recovered SNR of the test field CRGs indicates the reliability of the proposed self-supervised multistep deblending algorithm.

Figure 17 depicts the deblending performance of the 25th CRG. The unblended seismic record is shown in Fig. 17a as a reference for the deblending performance evaluation, and the corresponding pseudo-deblended record is shown in Fig. 17b. Figure 17c–f shows the deblended result and the attenuated blending noise using the proposed multistep algorithm and the MBNS, respectively. To better clarify the results, the local area marked by the red

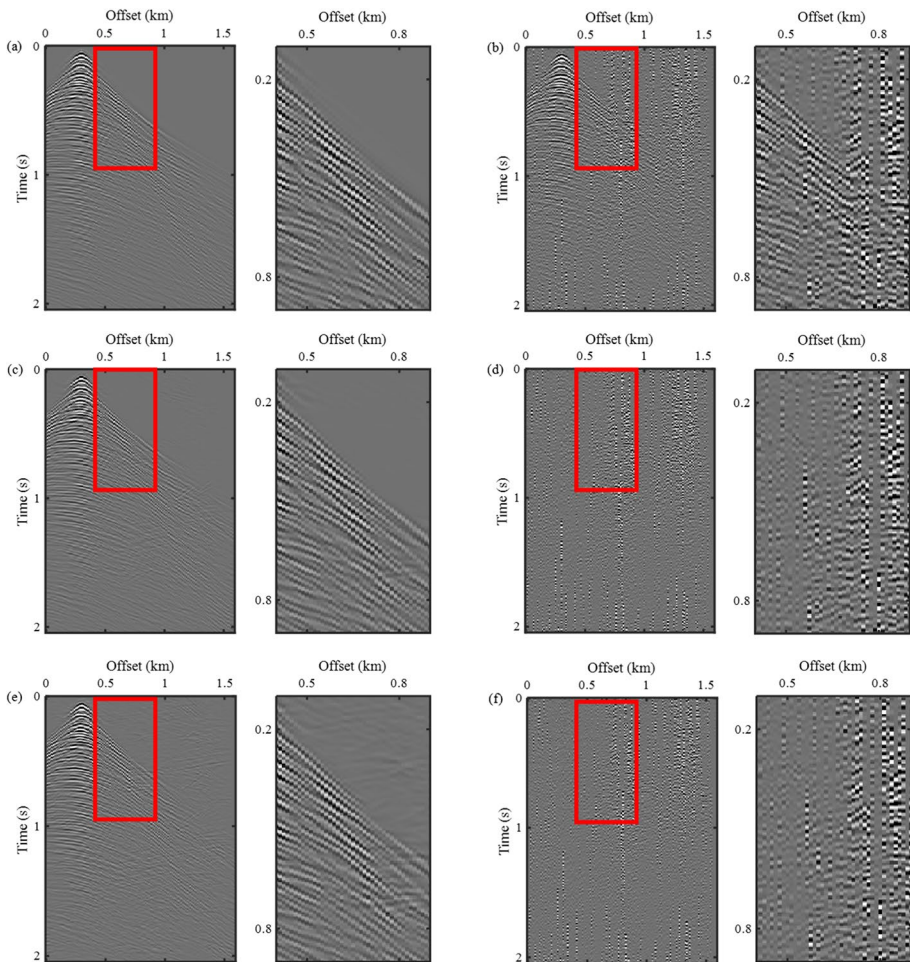


Fig. 17 Deblending performance comparison of the 25th CRG. **a** The reference unblended record; **b** The corresponding pseudo-deblended data; **c** Deblended result of the multistep (three-step) deblending algorithm; **d** The difference between **(b)** and **(c)**, i.e., the attenuated blending noise; **e** Deblended result of the MBNS (three-level) with three iterations; **f** The difference between **(b)** and **(e)**, i.e., the attenuated blending noise. The local area marked by the red box in each subpanel is enlarged and displayed to the corresponding right part

box in each subpanel is enlarged to the corresponding right side. The recovered SNRs are 11.90 dB and 10.79 dB for the deblended results obtained using the multistep (three-step) algorithm and the MBNS (three-level) with three iterations, respectively. In addition to the quantitative evidence of a higher recovered SNR, the proposed self-supervised multistep algorithm provides a visually better deblended result over the MBNS (three-level) with three iterations for field data deblending to demonstrate its superiority.

In addition to the evaluation based on recovered SNR, we further use a local similarity map (Chen and Fomel, 2015a) to assess the deblending performance. Figure 18 depicts the local similarity between the attenuated blending noise and the deblended result for the 25th CRG. The similarity map of the proposed multistep (three-step) algorithm (Fig. 18a) is lower than that of the MBNS (three-level) with three iterations (Fig. 18b), which indicates lower signal leakage and thus better deblending performance using the proposed multistep algorithm.

In order to further validate the self-supervised multistep deblending algorithm using an adaptive training set, the deblending performance of a specific CSG is displayed in Fig. 19 after deblending all CRGs. Figure 19a shows the unblended result of the 100th CSG as a reference to evaluate the deblending performance, and the corresponding pseudo-deblended record is shown in Fig. 19b with coherent blending noise contamination. Figure 19c–f displays the deblended result and the attenuated coherent blending noise using the multistep deblending algorithm and the MBNS, respectively. The recovered SNRs are 13.84 dB and 12.52 dB, respectively, and the proposed algorithm shows a quantitatively better deblending performance. For clear illustrations, we enlarge the areas where the blending noise dominates, marked by the red box in each subpanel, and display them to the corresponding right part. Figure 19c displays the deblended result obtained using the proposed method, which effectively attenuates the coherent blending noise, as shown in Fig. 19d. In contrast, there is some remaining coherent blending

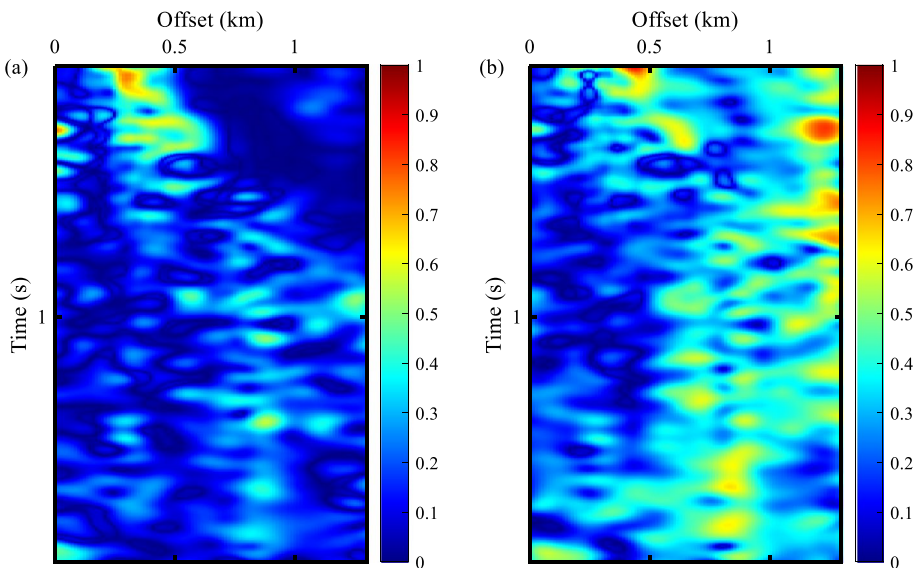


Fig. 18 Similarity maps between the deblended result and the attenuated blending noise of the 25th CRG using **a** the multistep (three-step) deblending algorithm and **b** the MBNS (three-level) with three iterations

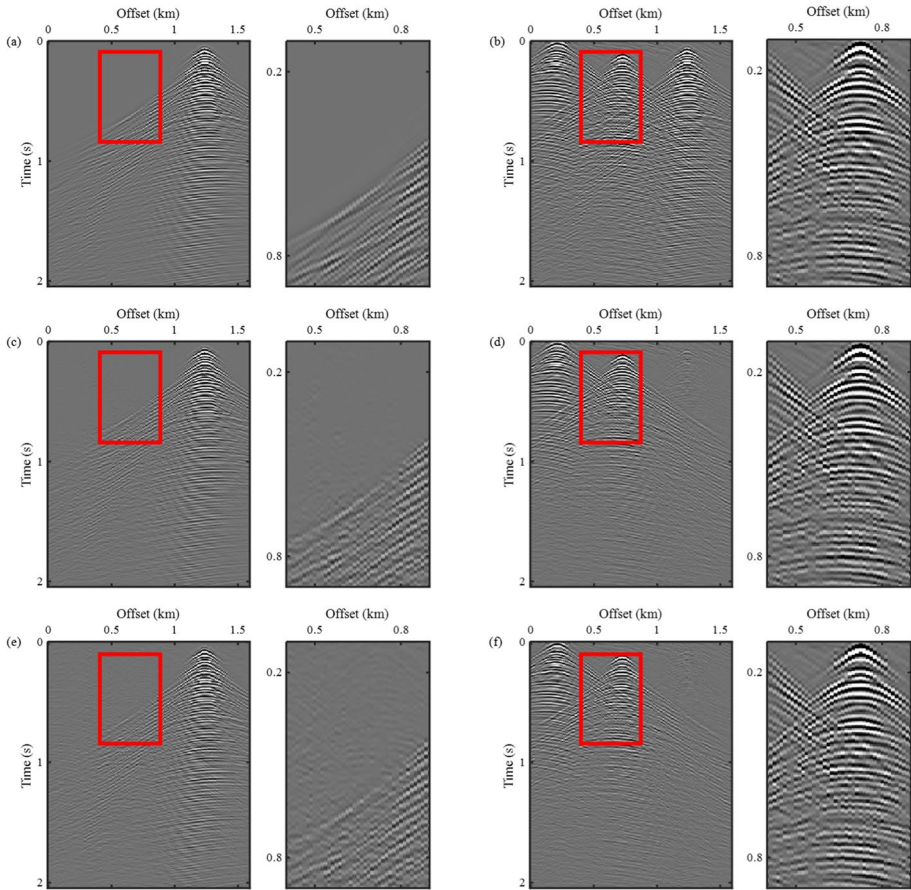


Fig. 19 Deblending performance comparison of the 100th CSG. **a** The reference unblended record; **b** The corresponding pseudo-deblended data; **c** Deblended result of the multistep (three-step) deblending algorithm; **d** The difference between **(b)** and **(c)**, i.e., the attenuated blending noise; **e** Deblended result of the MBNS (three-level) with three iterations; **f** The difference between **(b)** and **(e)**, i.e., the attenuated blending noise. The local area marked by the red box in each subpanel is enlarged and displayed to the corresponding right part

interference in the deblended result obtained using the MBNS (Fig. 19e), and the attenuated coherent blending noise is displayed in Fig. 19f.

To assess the signal preservation ability of different algorithms, Fig. 20 shows the similarity maps of the 100th CSG between the deblended result and the attenuated coherent blending noise using the proposed multistep deblending algorithm and the MBNS. The proposed algorithm generates a lower similarity map (Fig. 20a) almost everywhere when compared with the MBNS (Fig. 20b), which indicates the superiority of the proposed multistep deblending algorithm. Detailed comparisons demonstrate that the multistep algorithm is superior to the MBNS with similar computational cost during field data deblending.

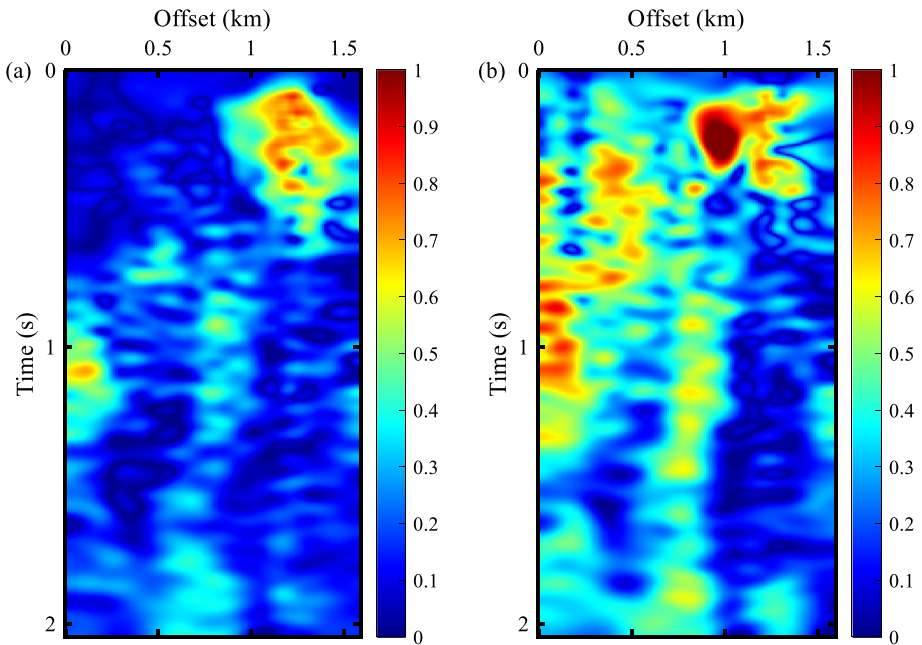


Fig. 20 Similarity maps between the debledned result and the attenuated blending noise of the 100th CSG using **a** the multistep (three-step) debrending algorithm and **b** the MBNS (three-level) with three iterations

4 Conclusions

We propose a self-supervised multistep debrending workflow to obtain debledned data for subsequent processing steps. In contrast to the supervised debrending algorithm, which requires clean labels, the self-supervised debrending algorithm uses the CSGs of pseudo-debledned data to adaptively construct the training sets. The CSGs are regarded as the labels, and the constructed results with the blending noise contamination are used as the training input. The multistep debrending strategy trains distinct individual networks to quantify the remaining blending noise in a flexible way. The predicted signal of the previous optimized network is used for blending noise estimation–subtraction to update the input of the current network. Besides, the optimized parameters of the previous network initialize the training of the current network based on transfer learning in order to maximize the training efficiency and stability. The SSIM is also incorporated flexibly into the loss function of subsequent network training for accurate remaining weak blending noise attenuation. Finally, the optimized individual networks are used sequentially to debrend the CRGs of the raw pseudo-debledned data. With CSG-assisted adaptive training data, the proposed multistep debrending strategy, without requiring clean labels, promotes the debrending performance. The proposed self-supervised multistep debrending algorithm is positively supported by synthetic and field data examples and achieves superior debrending performance when compared with the multilevel blending noise strategy in terms of weak signal preservation and blending noise attenuation.

Acknowledgements This research was financially supported by the CNPC Innovation Found (2023DQ02-0612), the National Natural Science Foundation of China (41874128) and the Fundamental Research Funds

for the Central Universities. We appreciate Jiakuo Li and Zhuowei Li at Tongji University for their helpful discussions. We also appreciate the Editor in Chief and reviewers for their critical and constructive comments to make the paper better.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

- Abma R, Manning T, Tanis M, Yu J, Foster M (2010) High quality separation of simultaneous sources by sparse inversion. In: Paper read at 72nd EAGE Conference and Exhibition
- Baardman R, Hegge R, Zwartjes P (2020) Deblending via supervised transfer learning-DSA field data example. In: Paper read at 82nd EAGE annual conference & exhibition
- Beasley CJ (2008) A new look at marine simultaneous sources. *Lead Edge* 27(7):914–917
- Berkhout A, Blacqui ere G, Verschuur E (2008) From simultaneous shooting to blended acquisition. In: Paper read at SEG annual meeting
- Chen Y (2015) Iterative deblending with multiple constraints based on shaping regularization. *IEEE Geosci Remote Sens Lett* 12(11):2247–2251
- Chen Y, Fomel SJG (2015a) Random noise attenuation using local signal-and-noise orthogonalization. *Geophysics* 80(6):WD1–WD9
- Chen Y, Fomel S, Hu J (2014a) Iterative deblending of simultaneous-source seismic data using seislet-domain shaping regularization. *Geophysics* 79(5):V179–V189
- Chen Y, Yuan J, Jin Z, Chen K, Zhang L (2014b) Deblending using normal moveout and median filtering in common-midpoint gathers. *J Geophys Eng* 11(4):045012
- Chen Y, Fomel S, Abma R (2023) Joint deblending and source time inversion. *Geophysics* 88(1):WA27–WA35
- Cheng J, Sacchi MD (2015b) A fast rank-reduction algorithm for 3D deblending via randomized QR decomposition. In: Paper read at SEG Annual Meeting
- Dai W, Wang X, Schuster GT (2011) Least-squares migration of multisource data with a deblurring filter. *Geophysics* 76(5):R135–R146
- Gan S, Wang S, Chen Y, Chen X, Xiang K (2016) Separation of simultaneous sources using a structural-oriented median filter in the flattened dimension: computers. *Geosciences* 86:46–54
- Hampson G, Stefani J, Herkenhoff F (2008) Acquisition using simultaneous sources. *Lead Edge* 27(7):918–923
- Han D, Wang B, Li J (2022) Consistent convolution kernel design for missing shots interpolation using an improved U-net. *Geophys Prospect* 70(7):1193–1211
- Huang W, Wang R, Gong X, Chen Y (2017) Iterative deblending of simultaneous-source seismic data with structuring median constraint. *IEEE Geosci Remote Sens Lett* 15(1):58–62
- Huang H, Wang T, Cheng J, Xiong Y, Wang C, Geng J (2022) Self-supervised deep learning to reconstruct seismic data with consecutively missing traces. *IEEE Trans Geosci Remote Sens* 60: 5911514
- Huo S, Luo Y, Kelamis PG (2012) Simultaneous sources separation via multidirectional vector-median filtering. *Geophysics* 77(4):V123–V131
- Ibrahim A, Sacchi MD (2014) Simultaneous source separation using a robust Radon transform. *Geophysics* 79(1):V1–V11
- Kontakis A, Verschuur D (2014) Deblending via sparsity-constrained inversion in the focal domain. In: Paper read at 76th EAGE conference and exhibition
- Li X, Wu B, Zhu X, Yang H (2021) Consecutively missing seismic data interpolation based on coordinate attention unet. *IEEE Geosci Remote Sens Lett* 19:3005005
- Lin R, Bahia B, Sacchi MD (2021) Iterative deblending of simultaneous-source seismic data via a robust singular spectrum analysis filter. *IEEE Trans Geosci Remote Sens* 60:5904110
- Lin R, Sacchi MD (2020) Separation of simultaneous sources via coherence pass robust radon operators. In: Paper read at SEG annual meeting
- Liu Y, Liu C, Wang D (2009) A 1D time-varying median filter for seismic random, spike-like noise elimination. *Geophysics* 74(1):V17–V24

- Liu N, Wu L, Wang J, Wu H, Gao J, Wang D (2022) Seismic data reconstruction via wavelet-based residual deep learning. *IEEE Trans Geosci Remote Sens* 60:4508213
- Mahdad A (2012) Deblending of seismic data: Ph.D thesis, TU Delft
- Mahdad A, Doulgeris P, Blacquièrre G (2011) Separation of blended data by iterative estimation and subtraction of blending interference noise. *Geophysics* 76(3):Q9–Q17
- Park MJ, Sacchi M (2020) Automatic velocity analysis using convolutional neural network and transfer learning. *Geophysics* 85(1):V33–V43
- Poole G, Stevens K, Maraschini M, Mensch T, Siliqi R (2014) Blended dual-source acquisition and processing of broadband data. In: Paper read at 76th EAGE conference and exhibition
- Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: Paper read at international conference on medical image computing and computer-assisted intervention.
- Sun J, Slang S, Elboth T, Larsen Greiner T, McDonald S, Gelius L-J (2020) A convolutional neural network approach to deblending seismic data. *Geophysics* 85(4):WA13–WA26
- Sun J, Hou S, Vinje V, Poole G, Gelius L-J (2022) Deep learning-based shot-domain seismic deblending. *Geophysics* 87(3):V215–V226
- Verschuur D, Berkhout A (2011) Seismic migration of blended shot records with surface-related multiple scattering. *Geophysics* 76(1):A7–A13
- Walker C, Hays D, Monk D (2013) Blended source single sided full azimuth ocean bottom seismic acquisition. In: Paper read at SEG annual meeting
- Wang B, Geng J (2019) Efficient deblending in the PFK domain based on compressive sensing. *IEEE Trans Geosci Remote Sens* 58(2):995–1003
- Wang K, Hu T (2022) Deblending of seismic data based on neural network trained in the CSG. *IEEE Trans Geosci Remote Sens* 60:5907712
- Wang B, Li J, Luo J, Wang Y, Geng J (2021a) Intelligent deblending of seismic data based on u-net and transfer learning. *IEEE Trans Geosci Remote Sens* 59(10):8885–8894
- Wang B, Li J, Han D (2022a) Iterative deblending using MultiResUNet with multilevel blending noise for training and transfer learning. *Geophysics* 87(3):V205–V214
- Wang B, Li J, Han D, Song J (2022b) Deblending and recovery of incomplete blended data via MultiResU-net. *Surv Geophys* 43(6):1901–1927
- Wang K, Hu T, Wang S (2022c) An unsupervised learning approach to deblend seismic data from denser shot coverage surveys. *Geophys J Int* 231(2):801–816
- Wang B, Chen X, Li J, Xiong D, Song J (2023) High dimensional multistep deblending using supervised training and transfer learning. *Geophysics* 88(1): WA149-WA159
- Wang S, Hu W, Yuan P, Wu X, Zhang Q, Nadukandi P, Ocampo Botero G, Chen J (2021b) Seismic deblending by self-supervised deep learning with a blind-trace network. In: Paper read at the first international meeting for applied geoscience & energy
- Xu W, Zhou Y, Liu D, Wang X, Chen W (2022) Seismic intelligent deblending via plug and play method with blended CSGs trained deep CNN Gaussian denoiser. *IEEE Trans Geosci Remote Sens* 60:5913413
- Xue Y, Chen Y, Jiang M, Duan H, Niu L, Chen C (2022) Unsupervised seismic data deblending based on the convolutional autoencoder regularization. *Acta Geophys* 70(3):1171–1182
- Yang L, Chen W, Liu W, Zha B, Zhu L (2020) Random noise attenuation based on residual convolutional neural network in seismic datasets. *IEEE Access* 8:30271–30286
- Yu J, Wu B (2022) Attention and hybrid loss guided deep learning for consecutively missing seismic data reconstruction. *IEEE Trans Geosci Remote Sens* 60:5902108
- Zhang Q, Mao W, Zhou H, Zhang H, Chen Y (2018) Hybrid-domain simultaneous-source full waveform inversion without crosstalk noise. *Geophys J Int* 215(3):1659–1681
- Zu S, Zhou H, Chen H, Zheng H, Chen Y (2017) Two field trials for deblending of simultaneous source surveys: why we failed and why we succeeded? *J Appl Geophys* 143:182–194
- Zu S, Cao J, Qu S, Chen Y (2020) Iterative deblending for simultaneous source data using the deep neural network Iterative deblending with DNN. *Geophysics* 85(2):V131–V141

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.