Petroleum Science 21 (2024) 1649–1659

Contents lists available at ScienceDirect

Petroleum Science

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

Original Paper

A hybrid WUDT-NAFnet for simultaneous source data deblending

Chao-Fan Ke [a, b, c], Shao-Huan Zu [a, b, c, *], Jun-Xing Cao [a, b, c], Xu-Dong Jiang [a, b, c], Chao Li [a, c], Xing-Ye Liu [a, c]

[a] Key Laboratory of Earth Exploration and Information Technology of Ministry of Education, Chengdu University of Technology, Chengdu 610059, Sichuan, China
[b] State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu 610059, Sichuan, China
[c] College of Geophysics, Chengdu University of Technology, Chengdu 610059, Sichuan, China

ARTICLE INFO

Article history:
Received 4 December 2022
Received in revised form
15 September 2023
Accepted 25 December 2023
Available online 29 December 2023

Edited by Jie Hao and Meng-Jiao Zhou

Keywords:
Simultaneous-source
Deblending
Deep learning
Transformer

ABSTRACT

Simultaneous source technology, which reduces seismic survey time and improves the quality of seismic data by firing more than one source with a narrow time interval, is compromised by the massive blended interference. Therefore, deblending algorithms have been developed to separate this interference. Recently, deep learning (DL) has been proved its great potential in suppressing the interference. The most popular DL method employs neural network as a filter to attenuate the blended noise in an iterative estimation and subtraction framework (IESF). However, there are still amplitude distortion and blended noise residual problems, especially when dealing with weak signal submerged in strong interference. To address these problems, we propose a hybrid WUDT-NAFnet, which contains two sub-networks. The first network is a wavelet based U-shape deblending transformer network (WUDTnet), incorporated into IESF as a robust regularization term to iteratively separate the blended interference. The second network is a nonlinear activate free network (NAFnet) designed to recover the event amplitude and further suppress the weak noise residual in IESF. With the hybrid network, the blended noise can be separated purposefully and accurately. Examples using synthetic and field seismic data demonstrate that the WUDT-NAFnet outperforms traditional curvelet transform (CT) based method and the deblending transformer (DT) model in terms of deblending. Additionally, for field applications, the data augmentation method of bicubic interpolation is applied to mitigate the feature difference between synthetic and field data. Consequently, the trained network exhibits strong signal preservation ability in numerical field example without requiring additional training.

1. Introduction

Traditional seismic acquisition requires a large time interval to prevent record overlap, while simultaneous source technology firing two or more sources to obtain densely sampled data, can greatly enhance the efficiency of seismic acquisition. Nevertheless, the benefit of simultaneous source acquisition is hindered by the intense blended interference (van Borselen et al., 2012; Mahdad et al., 2011). Direct imaging of simultaneous source data can be realized by the least-squares reverse time migration (LSRTM) (Xue et al., 2016; Li et al., 2017). However, achieving accurate direct imaging becomes challenging without precise velocity information. To give full play to the economic advantage of simultaneous source technology, an effective deblending algorithm is still urgently needed.

Simultaneous source technology usually adopts dithering scheme to generate coherency difference in common receiver (CRG) or offset gather (COG). With the dithering time, simultaneous source data in CRG can be sorted into the pseudo-deblended record where the component of the reference source is coherent while that of the other sources are incoherent. Based on this characteristic, many deblending algorithms are developed, which can be summarized into two categories roughly. The first category is similar to denoising, directly attenuating the crosstalk noise with high efficiency. Wapenaar et al. (2012) implemented a direct matrix inversion method by inserting a filter to the least-squares inversion algorithm. Sun et al. (2020) used numerically blended field data to train a convolutional neural network (CNN), which was viewed as a

* Corresponding author.
E-mail address: zushaohuan19@cdut.edu.cn (S.-H. Zu).

https://doi.org/10.1016/j.petsci.2023.12.017

filter to denoise the incoherent interference. Nakayama and Blacquière (2021) proposed a U-net network based on residual block to implement deblending, trace reconstruction, and low-frequency extrapolation simultaneously. The second category views deblending as an iterative estimation problem, which enjoys high deblending performance. Mahdad et al. (2011) proposed an iterative estimation and subtraction framework (IESF), and used $f-k$ filter with thresholding to gradually recover the signal. Since the incoherent noise and trace missing increase the rank of Hankel constructed from the monochromatic frequency, Cheng and Sacchi (2015) introduced an iterative rank reduction method based on singular spectrum analysis to separate and recover the simultaneous source data. Moreover, by formulating the deblending problem as a regularization problem, the conventional optimization methods, e.g., projection onto convex set (POCS) algorithm (Zhou, 2017), and iterative shrinkage thresholding algorithm (ISTA) can be introduced to separate the incoherent interference (Qu et al., 2016). Based on the compressive sensing principle, some researchers introduce the sparse constraint as the regularization term to the deblending inversion problem. Chen et al. (2014) proposed to iteratively suppress the blended interference using shaping regularization term in the seislet domain. Gan et al. (2016) introduced separate blended data in common shot gathers (CSGs) with a novel iterative seislet-frame thresholding approach. Owning to the fact that the incoherent interference affects the accuracy of estimated slope, reducing the sparsity of seislet transform, Zhou et al. (2018) implemented a seislet transform with the velocity-slope conversion scheme to deal with the blended noise. Since the coherent signal and incoherent noise have different energy distribution and values in the Radon domain, Zhang et al. (2015) proposed to solve deblending in common midpoint domain by using high-resolution Radon transform. Xue et al. (2017) used a high-order Radon transform to iteratively suppress the blended interference. Lin and Sacchi (2020) adopted robust sparse Radon transform with coherency-pass operator to separate the blended data. Due to the curvelet coefficient difference between the coherent signal and the blended interference, Zhang et al. (2021) implemented a non-equispaced curvelet transform (CT) to suppress the blended interference in the non-uniformly sampled simultaneous source data.

In recent years, deep learning (DL) has been widely used to iteratively deal with simultaneous-source data. Based on the inversion principle, Zu et al. (2020) embedded the trained network as a regularization term into IESF to optimize the deblending performance. Wang et al. (2021) used a U-shape network to iteratively separate the blended noise and proposed transfer learning for field data application. Wang and Hu (2021) manually simulated the blending common shot gathers (CSGs) as training data, which can adapt to the given survey. Xu et al. (2022a,b) embedded a DnCNN network with real spectral normalization into alternating direction method of multipliers (ADMM) framework and trained the network using the manually blending field CSGs. For better adapting IESF, Wang et al. (2022a,b,c) presented a multiresolution ResU-net and trained it using multi-level blended noise. For improving the feature extraction ability of transformer model to the blended data, Zu et al. (2022) proposed a deblending transformer (DT) network with horizontal, vertical and local self-attention mechanism. In order to make the trained network fully adopt to the given survey, Sun et al. (2022) manually blended the unblended shot gathers, which were gathered at the end to each sail line. Xu et al. (2022a,b) proposed to solve the deblending problem by manually blending the CSGs and using them to train a convolutional autoencoder, which achieved better deblending performance than traditional training scheme. To make CNN applicable to the field data, Wang et al. (2022a,b,c) developed a data augmentation method, which used shuffled deblending noise as the input and the inaccurate labels as the output to train the network. To improve deblending performance and reconstruction accuracy, Wang et al. (2022a,b,c) designed a novel workflow, which firstly trained the designed MultiResUnet using multi-level blended noise in the common receiver domain for iteratively joint deblending and trace reconstruction, then the training data was used again to fine tune the trained network for pure missing trace reconstruction in the common shot domain. Wang et al. (2023) proposed to train the neural network with the blended data, data with low amplitude blended noise, and unblended data, which can improve network performance when iteratively deblending. The previous descried DL based methods, can suppress the blended interference by iterative framework to some extent, however, the problem of amplitude distortion and weak noise leakage still exists.

In this article, we first introduce a wavelet based U-shape deblending transformer model called WUDTnet. WUDTnet is a multi-level encoder-decoder transformer designed to efficiently capture local and global representations of the simultaneous source data. It employs discrete Haar wavelet transform (DHWT) for down-sampling and inverse Haar wavelet transform (IHWT) for up-sampling. To accelerate the convergence, an improved DT block is introduced, which includes additional fusion layer and normalization layers at different sections. The proposed WUDTnet is considered as the strong denoiser and embedded in IESF to iteratively separate the blended noise. Next, we apply the nonlinear activation free network (NAFnet) to recover the signal with the weak incoherent residual from the deblended results of IESF, which adapts to the amplitude range of seismic data by replacing the nonlinear activation function with a Simple gate module. Experimental results on one synthetic and two field examples demonstrate that the presented WUDT-NAFnet achieves superior deblending performance and higher signal fidelity compared to both the traditional CT based method and the DT model.

2. Theory

2.1. IESF

Taking a simultaneous source survey with two sources as an example. In the field, one source is typically designated as the reference source and does not have delay time. The second source, on the other hand, fires with a dithering time of Γ_2. Therefore, the blended common receiver gather (CRG) $\mathbf{b}_{\text{ble}}$ is viewed as the pseudo-deblended data of the reference source. The blending process is expressed as

$$\mathbf{b}_{\text{ble}} = \mathbf{b}_1 + \Gamma_2 \mathbf{b}_2 \tag{1}$$

where $\mathbf{b}_i$ represents the record of the ith source. Applying Γ_2^{-1} to Eq. (1), we can obtain the pseudo-deblended record of the other source:

$$\Gamma_2^{-1}\mathbf{b}_{\text{ble}} = \Gamma_2^{-1}\mathbf{b}_1 + \mathbf{b}_2. \tag{2}$$

Combining Eqs. (1) and (2), the brief equation can be obtained:

$$\mathbf{B} = \mathbf{F}\mathbf{D} \tag{3}$$

where

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_{\text{ble}} \\ \Gamma_2^{-1}\mathbf{b}_{\text{ble}} \end{bmatrix}, \mathbf{F} = \begin{bmatrix} \mathbf{I} & \Gamma_2 \\ \Gamma_2^{-1} & \mathbf{I} \end{bmatrix}, \mathbf{D} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}. \tag{4}$$

Considering simultaneous source deblending as a denoising task, the recovered signal $\widehat{\mathbf{D}}$ obtained by using DL method can be

formulated as

$$\widehat{\mathbf{D}} = \mathscr{T}(\mathbf{B}; \Theta_{\text{trained}}) \tag{5}$$

where $\mathscr{T}(\cdot)$ represents the designed neural network, Θ_{trained} is the trained network parameter. Directly deblending using Eq. (5) is highly efficient, but may damages the coherent signal. Therefore, Zu et al. (2020) embedded the trained network into IESF and demonstrated the advantage of the iterative scheme, especially in protecting the weak signal. Combining Eqs. (3) and (5), the IESF can be expressed as

$$\widehat{\mathbf{D}}_{i+1} = \mathscr{T}(\mathbf{B} - (\mathbf{F} - \mathbf{I})\widehat{\mathbf{D}}_i; \Theta_{\text{trained}}) \tag{6}$$

where $\widehat{\mathbf{D}}_{i+1}$ denotes the recovered result in the $(i+1)$th iteration. In this framework, there are two steps to constrain the incoherent interference. The first step uses the trained network $\mathscr{T}(\cdot; \Theta_{\text{trained}})$ to strictly suppress the incoherent interference and preserve the coherent signal. The second step applies the recovered coherent signal $\widehat{\mathbf{D}}_i$ and blending operator $\mathbf{F}$ to estimate the blended noise and separate it from the incoherent record $\mathbf{B}$. Compared to denoising task, the IESF scheme provides greater flexibility in separating blended data. This is because denoising tasks face challenges in suppressing blended noise without causing damage to the coherent signal. In the IESF scheme, the first step only requires to pass the coherent signal, allowing for some damage. The second step focuses on estimating and accurately suppressing the incoherent interference.

2.2. Hybrid WUDT-NAFnet

In this section, we introduce a WUDTnet architecture, incorporating an improved DT block, a DHWT down-sampling layer, and an IHWT up-sampling layer. Subsequently, we embed WUDTnet into the IESF to gradually recover the coherent signal. To address issues such as amplitude distortion and weak noise leakage, we employ a NAFnet for the final data reconstruction (Chen et al., 2022). The general architecture of our hybrid WUDT-NAFnet is illustrated in Fig. 1.

2.2.1. WUDTnet

The transformer model has demonstrated great potential in handling image denoising tasks. However, the efficiency of denoising is compromised by the high computational cost associated with the self-attention layer. Therefore, Zu et al. (2022) proposed a DT model with feature split based self-attention mechanism to alleviate this problem and achieve good deblending performance. The DT model, characterized by its shallow architecture, performs well when increasing the depth or the hidden dimension of features. However, this advantage is hindered by the fact that the cost of memory grows quadratically with the increase in network depth or width. Although using gradient accumulation technology or upgrading equipment can address this problem to some extent, for a better balance between deblending performance and computational cost, we propose the WUDTnet. The detail architecture of the designed network is illustrated in Fig. 1(a). Given a blended data $\mathbf{x} \in \mathbb{R}^{H \times W \times 1}$ whose size is $H \times W$, the WUDTnet first applies an overlap patch embedding layer to obtain shallow hidden features $\mathbf{x}_0 \in \mathbb{R}^{H \times W \times C}$, where C denotes the number of hidden dimensions. Next, these features pass through a multi-level U-shape architecture. Each level contains multiple improved DT blocks. Moreover, for efficient features down sampling and up sampling, we apply DHWT and IHWT, respectively. To flexibly fuse the low-level features and high-level features, the encoder and decoder are concatenated via skip connection.

2.2.2. Improved DT block

For computational efficiency, Zu et al. (2022) combined a special feature split method with self-attention mechanism, which extracts global representations horizontally, vertically and locally. The self-attention mechanism can be expressed as follows:

$$[\mathbf{Q}, \mathbf{K}, \mathbf{V}] = Linear(\mathbf{z}_t) \tag{7}$$

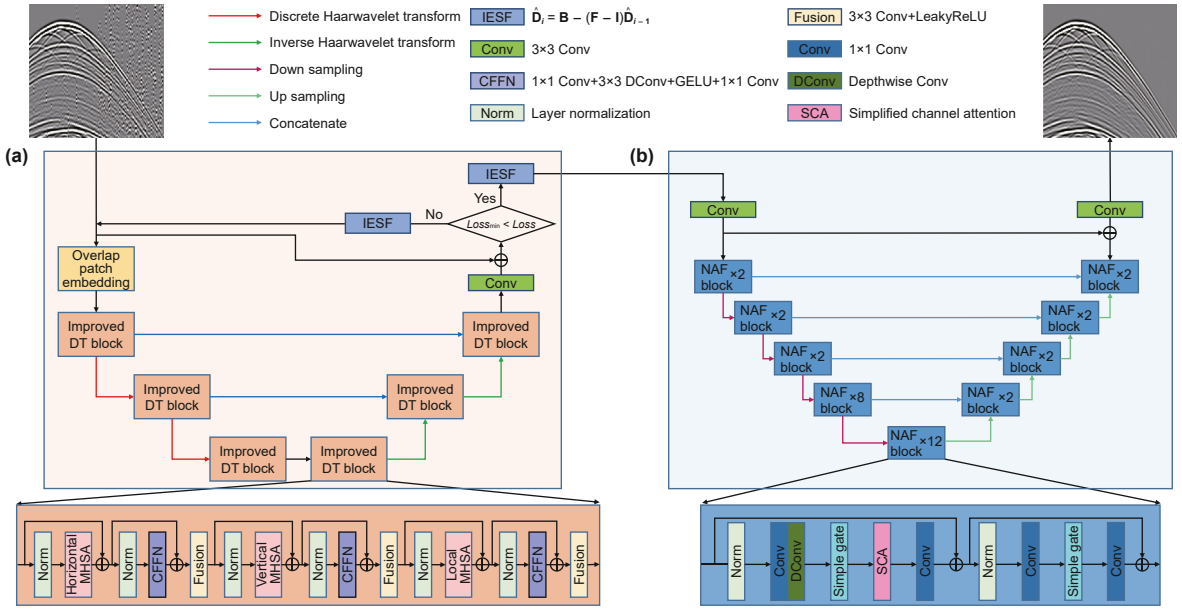

Fig. 1. The architecture of WUDT-NAFnet. It contains two sub-networks: **(a)** the designed WUDTnet, where *Loss* can be represented by Eq. (14), and **(b)** NAFnet.

$$\mathbf{W} = Softmax\left(\frac{\mathbf{Q}\mathbf{K}^{\mathrm{T}}}{\sqrt{D}}\right) \tag{8}$$

$$Atten(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathbf{W}\ \mathbf{V} \tag{9}$$

where $\mathbf{Q}$, $\mathbf{K}$ and $\mathbf{V}$ denotes the query, key and value sets, respectively. Given an input $\mathbf{z}_t \in \mathbb{R}^{N\times D}$ in the self-attention layer, $[\mathbf{Q},\mathbf{K},\mathbf{V}] \in \mathbb{R}^{N\times 3D}$ can be obtained through linear layer, where N is the length of the sequence and D is the dimension of the sequence. $\mathbf{W} \in \mathbb{R}^{N\times N}$ denotes the attention feature map, where $\mathbf{W}_{i,j}$ are based on the pairwise correlation between the query $\mathbf{Q}_i$ and key $\mathbf{K}_j$ representations, and $1/\sqrt{D}$ is applied to normalize the gradient. Finally, the attentioned $\mathbf{z}_t' \in \mathbb{R}^{N\times D}$ can be computed by the matrix product of attention map $\mathbf{W}$ and value set $\mathbf{V}$. Furthermore, the multi-head self-attention (MHSA) can be achieved by dividing $\mathbf{Q}$, $\mathbf{K}$ and $\mathbf{V}$ into M groups. Each group of $\mathbf{Q}$, $\mathbf{K}$ and $\mathbf{V}$ extracts different representations in parallel. In the improved DT block, the input $\mathbf{z} \in \mathbb{R}^{H\times W\times C}$ of feature split based MHSA module is transformed to $\mathbf{z}_t \in \mathbb{R}^{N\times D}$, where $t \in (h,v,l)$. For horizontal MHSA, we split the feature into $\mathbf{z}_h \in \mathbb{R}^{N\times(1\cdot W\cdot C)}$, where $(1\cdot W)$ denotes the horizontally extracted window and $N = H\cdot W/(1\cdot W)$. For vertical MHSA, we split the feature into $\mathbf{z}_v \in \mathbb{R}^{N\times(H\cdot 1\cdot C)}$, where $(H\cdot 1)$ denotes the vertically extracted window and $N = H\cdot W/(H\cdot 1)$. For local MHSA, we split the feature into $\mathbf{z}_l \in \mathbb{R}^{N\times(L\cdot L\cdot C)}$, where $(L\cdot L)$ denotes the locally extracted window whose size is L and $N = H\cdot W/(L\cdot L)$. These special features split scheme can efficiently balance the computational cost and deblending performance.

To further explore the advantages of the transformer model and expedite the network convergence, we develop an improved DT block. As shown in Fig. 1(a), each improved DT block contains three successive transformer layers with different self-attention mechanism. Before feeding into feature split based MHSA module, a layer normalization operation is used to alleviate the vanishing gradients problem and stable the network training. Moreover, to enhance efficiency and enable flexible feature fusion, the convolutional feed-forward network (CFFN) module is applied, which consists of four parts: a 1×1 convolution layer with $4C$ kernels, a 3×3 depth-wise convolution layer with $4C$ kernels, Gaussian error linear unit (GELU) activation function, and a 1×1 convolution layer with C kernels. Besides, we employ a residual connection between the Norm + MHSA layer and Norm + CFFN layer (He et al., 2016). Finally, a 3×3 convolution layer with LeakyReLU activation is implemented to fuse the attention feature after every transformer layer.

2.2.3. DHWT

The DHWT in WUDTnet is closely related with the down sampling operation. Traditional down sampling layer, for instance, pooling layer, and convolutional down sampling layer, only retains the part of information in the original image. In contrast, DHWT considers the characteristics of frequency and location, making the WUDTnet can correctly recover the detailed textures. Assuming $\mathbf{A}$ is a $H\times W$ matrix, where H and W are even. The 2D DHWT can be realized by:

$$\mathbf{E} = \mathbf{T}_{\mathrm{H}}\mathbf{A}\mathbf{T}_{\mathrm{W}}^{\mathrm{T}} = \begin{bmatrix} \mathscr{A} & \mathscr{H} \\ \mathscr{V} & \mathscr{D} \end{bmatrix} \tag{10}$$

where $\mathbf{E}$ is the transformed feature, $\mathbf{T}_{\mathrm{H}}$ denotes the column-wise DHWT, $\mathbf{T}_{\mathrm{W}}^{\mathrm{T}}$ denotes row-wise DHWT, and $\mathscr{A}$, $\mathscr{V}$, $\mathscr{H}$ and $\mathscr{D}$ are approximate matrices, vertical details, horizontal details and diagonal details, respectively. Defining $a_{i,j}$ is the element in the ith row and the jth column of $\mathbf{A}$, $\eta_{i,j}$, $\beta_{i,j}$, $\gamma_{i,j}$ and $\delta_{i,j}$ are the element of $\mathscr{A}$, $\mathscr{V}$, $\mathscr{H}$ and $\mathscr{D}$, respectively, the expression is as follows:

$$\begin{aligned} \eta_{i,j} &= \frac{a_{2i-1,2j-1} + a_{2i-1,2j} + a_{2i,2j-1} + a_{2i,2j}}{2} \\ \beta_{i,j} &= \frac{-a_{2i-1,2j-1} + a_{2i-1,2j} - a_{2i,2j-1} + a_{2i,2j}}{2} \\ \gamma_{i,j} &= \frac{-a_{2i-1,2j-1} - a_{2i-1,2j} + a_{2i,2j-1} + a_{2i,2j}}{2} \\ \delta_{i,j} &= \frac{a_{2i-1,2j-1} - a_{2i-1,2j} - a_{2i,2j-1} + a_{2i,2j}}{2} \end{aligned} \tag{11}$$

2.2.4. NAFnet

The NAFnet is utilized for preserving more coherent signal and suppressing the weak blended residual (Chen et al., 2022). As shown in Fig. 1(b), NAFnet is a multi-scale U-shape architecture with convolution down sampling layer and PixelShuffle up sampling layer (Shi et al., 2016). The encoder and decoder of NAFnet are based on NAFblock, which mainly consists of layer normalization, convolutional operator, Simple gate module, and simplified channel attention (SCA) module. The Simple gate can be realized by an element-wise multiplication:

$$Simple\ \ gate(\mathbf{x}_{o1},\mathbf{x}_{o2}) = \mathbf{x}_{o1}\odot\mathbf{x}_{o2} \tag{12}$$

where $\mathbf{x}_{o1} \in \mathbb{R}^{H\times W\times C/2}$ and $\mathbf{x}_{o2} \in \mathbb{R}^{H\times W\times C/2}$ are the two different parts of $\mathbf{x}_o \in \mathbb{R}^{H\times W\times C}$, and $\odot$ denotes element-wise multiplication. Furthermore, the SCA module is introduced into NAFblock to capture the global information and enhance computational efficiency:

$$SCA(\mathbf{x}) = \mathbf{x}^{*}pool(\mathbf{x}) \tag{13}$$

where *pool* denotes the adaptive average pooling operation, which extracts a scale factor of each channel, and * represents the channel-wise product operation.

3. Experiments

In this section, we assess the effectiveness of the proposed WUDT-NAFnet using synthetic and field data. Additionally, we compare the deblending performance by embedding the DT model and the CT-based method into IESF. The number of iterations in IESF plays a significant role in suppressing blended interference. For a fair comparison, we set the number of iterations in IESF for WUDTnet to 19, DT to 20, and the CT method to 40. Moreover, we use Eq. (14) to determine the best-recovered result in this range to ensure reliability, where L_2 represents mean-square error. Specially, to quantitatively and qualitatively analysis the contribution of each sub-network, we also illustrate the deblended results of WUDTnet, which is embedded in IESF.

$$Loss = L_2(\mathbf{B},\mathbf{F}\widehat{\mathbf{D}}_i) \tag{14}$$

To quantitatively assess the signal recovery capability of different algorithms, the metrics of the signal-to-noise ratio (SNR) is applied in all the following experiments:

$$\mathrm{SNR(dB)} = 10\log_{10}\frac{\|\mathbf{y}\|^2}{\|\mathbf{y}-\widehat{\mathbf{y}}\|^2} \tag{15}$$

where $\mathbf{y}$ represents the ground-truth seismic record and $\widehat{\mathbf{y}}$ denotes the estimated result. A higher SNR, indicates better deblending performance.

3.1. Implementation details

3.1.1. Training set

DL is a data-driven method where the richness of training samples is vital for obtaining effective neural network parameters. Our training set contains 673 synthetic CRGs and 400 raw field CRGs. The validation set contains 117 synthetic CRGs and 99 field CRGs. In addition, for field data application, 63 synthetic CRGs are resized from 1024×512 to 1500×300 by bicubic interpolation algorithm to simulate the low horizontal resolution feature of field data, which shortens the characteristic difference between the synthetic and real acquisition data. Moreover, before feeding the network, the unblended CRG is manually blended with a delay time of [−0.8 s, 0.8 s], and then cropped into patches, whose size are 64×64.

3.1.2. Loss function

As a common denoising task, we optimize the WUDTnet with L_2 loss, which is expressed as

$$L_2(\Theta) = \|\mathscr{F}(\mathbf{X};\Theta) - \mathbf{Y}\|_F^2, \tag{16}$$

where $\mathbf{X}$ stands for the input of the network, $\mathscr{F}(\cdot)$ denotes the adopted network architecture, Θ is the corresponding parameters to be updated, and $\mathbf{Y}$ represents the ground-truth.

However, L_2 loss may cause the problem that the predicted events are too continuous. In order to alleviate this matter and adapt the real seismic events, we apply Charbonnier loss function with $\varepsilon = 1e^{-6}$ to optimize the NAFnet, which can handle outliers and improve reconstructing performance:

$$Char(\Theta) = \sqrt{\|\mathscr{F}(\mathbf{X};\Theta) - \mathbf{Y}\|^2 + \varepsilon^2}. \tag{17}$$

3.1.3. Parameters

The WUDTnet employs a three level U-shape encoder-decoder. From the first level to the third level, the number of attention heads is [2, 4, 8], and the number of channels is [64, 128, 256]. For NAFnet, we implement a four level U-shape encoder-decoder, the number of channels is [32, 64, 128, 256].

The training procedure of the WUDT-NAFnet is shown in Table 1. We individually optimize the WUDTnet and NAFnet. The WUDTnet is firstly trained with blended inputs $\mathbf{B}_d$ and corresponding labels $\mathbf{D}_d$. Next, we start to update the NAFnet parameters, the blended inputs $\mathbf{B}_d$ are recovered by the trained WUDTnet with IESF. The iteration of IESF stops when L_2 loss reaches its minimum. Finally, NAFnet will be trained by the output of IESF $\widehat{\mathbf{D}}_d{}'$ and corresponding ground-truth $\mathbf{D}_d$.

Table 1
The implemented training procedure of the proposed method.

Input: Blended seismic data $\mathbf{B}_d$ and corresponding labels $\mathbf{D}_d$, delay times $\mathbf{F}_d$, WUDTnet model with parameters Θ_W, NAFnet model with parameters Θ_N, the epoch number of WUDTnet M_W^{epoch}, the epoch number of NAFnet M_N^{epoch}
1: Initialize parameters Θ_W and Θ_N under uniform distribution
2: for $i = 1, 2, \cdots, M_W^{epoch}$ do
3: Update Θ_W according to Eq. (16) with pairs of $\mathbf{B}_d$ and $\mathbf{D}_d$
4: end for 5: for $j = 1, 2, \cdots, M_N^{epoch}$ do
6: Initialize $L_{min} = 1.0$, $k = 1$
7: $\widehat{\mathbf{D}}_d^{(1)} \leftarrow \mathscr{F}(\mathbf{B}_d;\Theta_W)$, $L_{cur} \leftarrow L_2(\mathbf{B}_d, \mathbf{F}_d\widehat{\mathbf{D}}_d^{(1)})$
8: while $L_{min} > L_{cur}$ do
9: $k \leftarrow k+1$, $L_{min} \leftarrow min(L_{min}, L_{cur})$,
10: $\widehat{\mathbf{D}}_d^{(k)} \leftarrow \mathscr{F}(\mathbf{B}_d - (\mathbf{F}_d - \mathbf{I})\widehat{\mathbf{D}}_d^{(k-1)};\Theta_W)$, $L_{cur} \leftarrow L_2(\mathbf{B}_d, \mathbf{F}_d\widehat{\mathbf{D}}_d^{(k)})$
11: end while
12: $\widehat{\mathbf{D}}_d{}' \leftarrow \mathbf{B}_d - (\mathbf{F}_d - \mathbf{I})\widehat{\mathbf{D}}_d^{(k-1)}$
15: Update Θ_N according to Eq. (17) with pairs of $\widehat{\mathbf{D}}_d{}'$ and $\mathbf{D}_d$
16: end for 17: Output: Optimized parameters Θ_W and Θ_N

The networks used in the experiment are all trained by AdamW optimizer, where $\beta_1 = 0.9$, $\beta_2 = 0.999$, the weight decay rate is set to $1e^{-3}$, and the batch size is 264. The WUDTnet and NAFnet are separately optimized by L_2 and *Char* loss function for 100 epochs with the basic learning rate $5e^{-4}$ gradually reduced to $2.5e^{-4}$ with the cosine learning rate scheme. For DT model, we optimize it by L_2 loss function for 200 epochs with the basic learning rate $5e^{-4}$ gradually reduced to $1e^{-8}$. The network training is implemented on a personal workstation with an Intel(R) Xeon(R) Platinum 8180 processor and four Quadro GP100 GPUs. The training of DT, WUDTnet and NAFnet costs 73.62, 57.98 and 49.66 h, respectively. In addition, we spend 4.68 h to get the output of IESF $\widehat{\mathbf{D}}_d{}'$ using the trained WUDTnet. Therefore, the total training computation cost of our proposed algorithm is 112.32 h. The learning and validation curve of DT, WUDTnet and NAFnet are shown in Fig. 2(a)–(c), respectively, where the black line represents the learning curve,

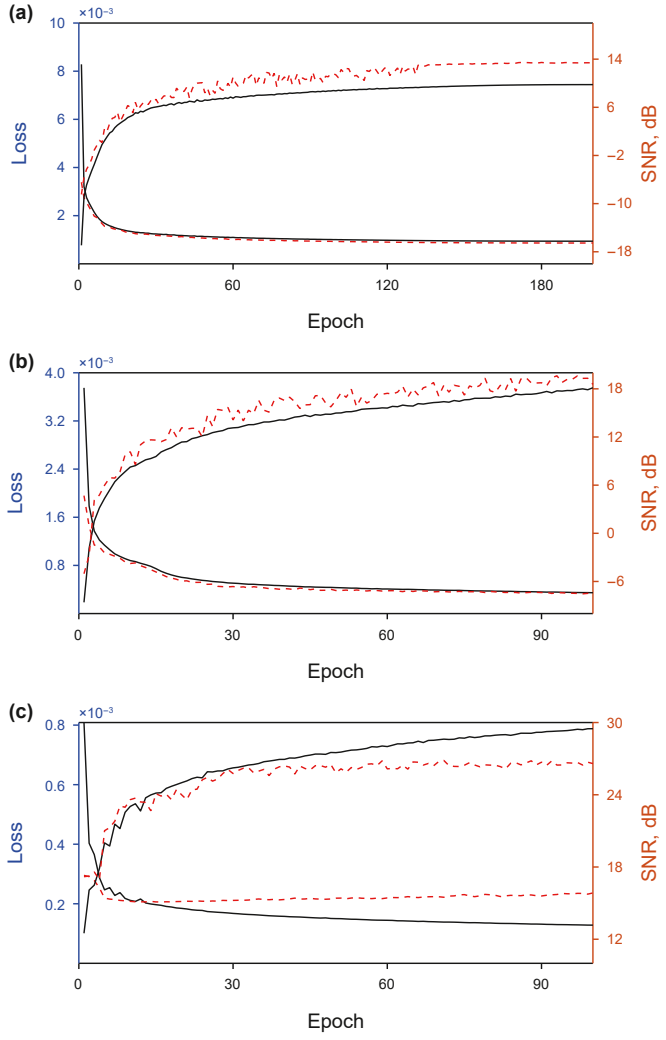

Fig. 2. The learning and validation curve of **(a)** DT, **(b)** WUDTnet and **(c)** NAFnet, respectively, where the black line represents the learning curve, and the red dashed represents the validation curve.

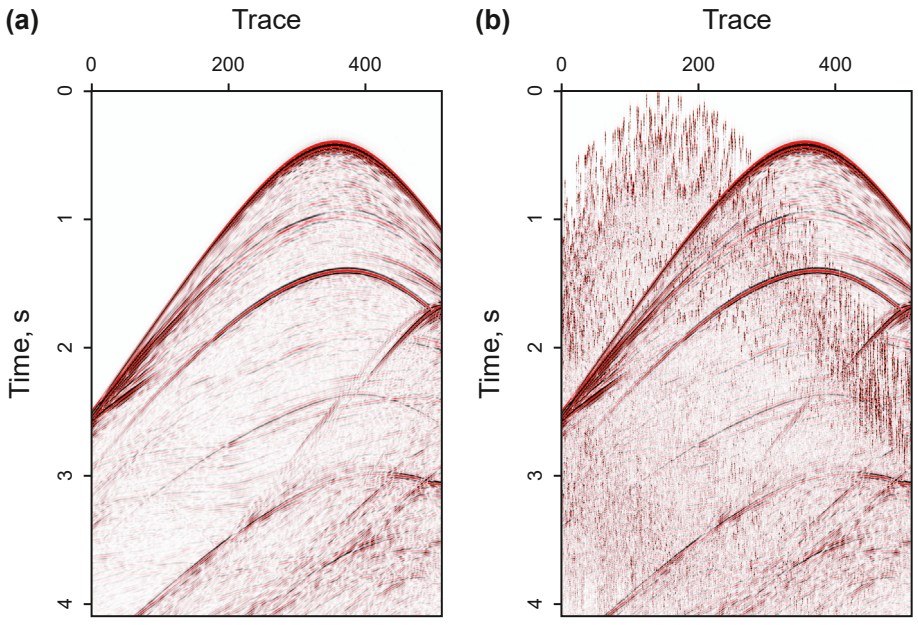

Fig. 3. Complex synthetic CRG example for comparison experiment. **(a)** Unblended CRG. **(b)** Pseudo-deblended CRG.

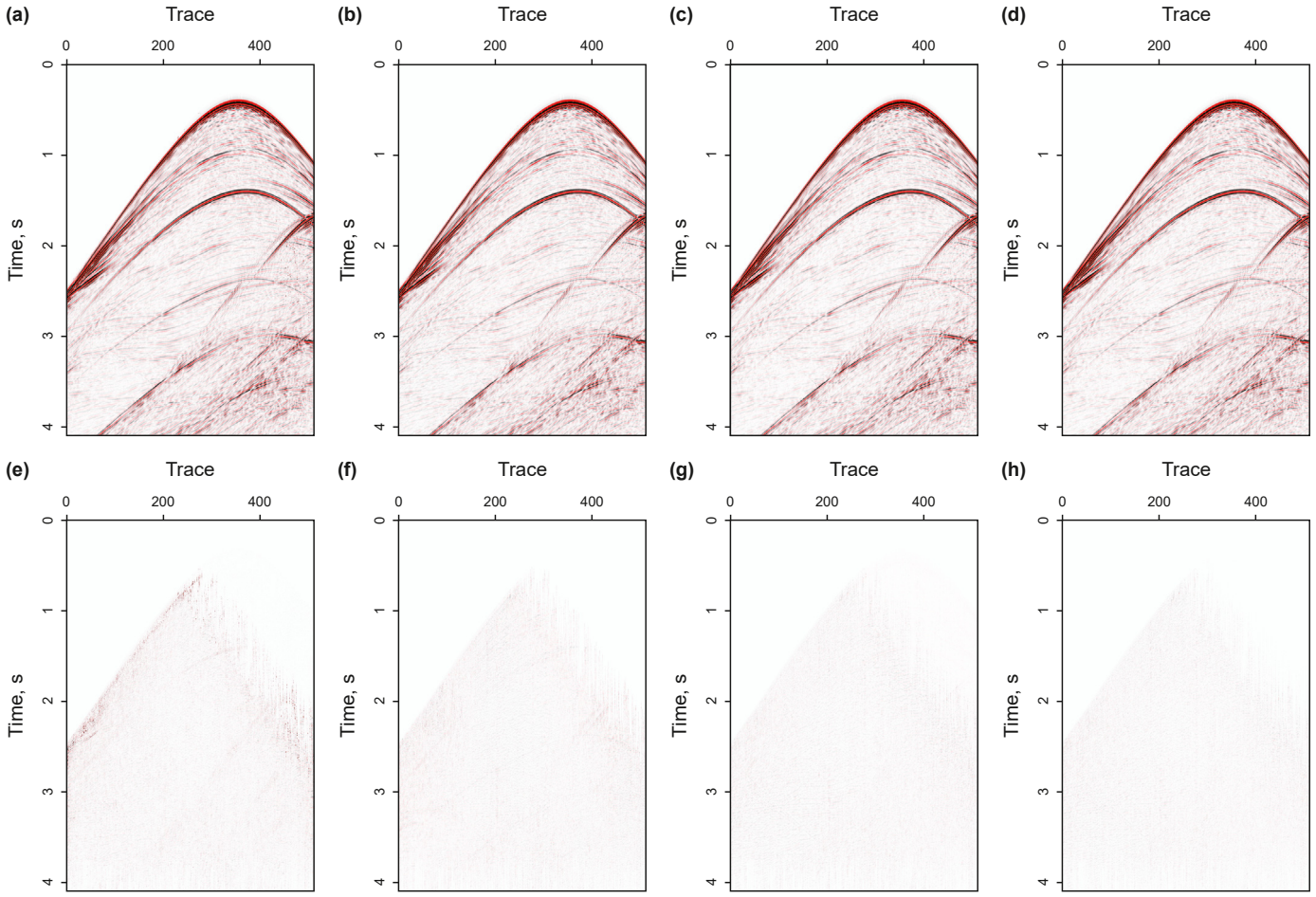

Fig. 4. Deblended results using **(a)** the CT based method, **(b)** the DT model, **(c)** the WUDTnet, and **(d)** the presented method. Deblended errors corresponding to **(e)** the CT based method, **(f)** the DT model, **(g)** the WUDTnet, and **(h)** the presented method.

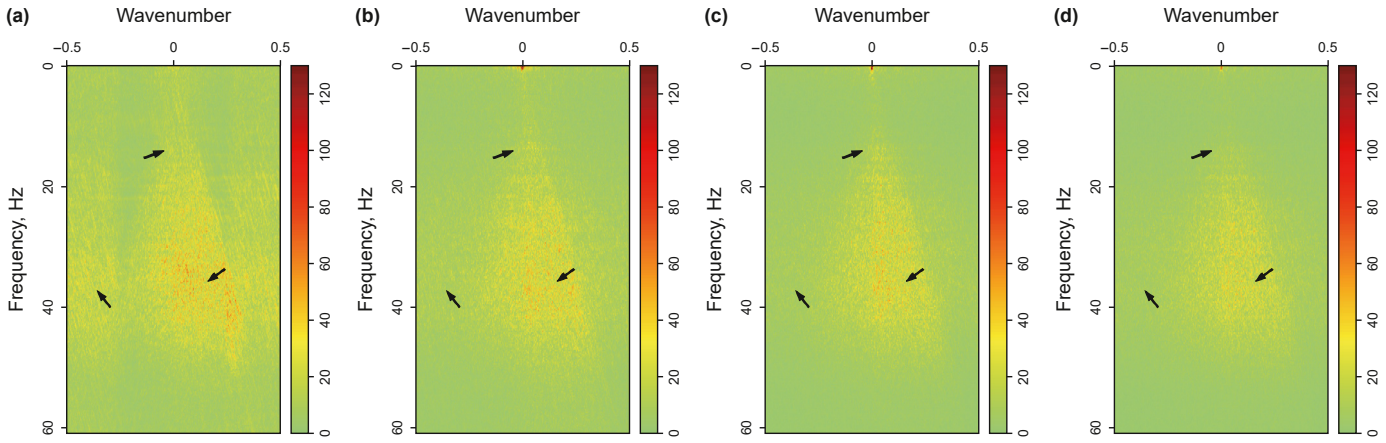

Fig. 5. Deblended errors $f-k$ spectra comparison of **(a)** the CT based method, **(b)** the DT method, **(c)** the WUDTnet model and **(d)** the proposed method.

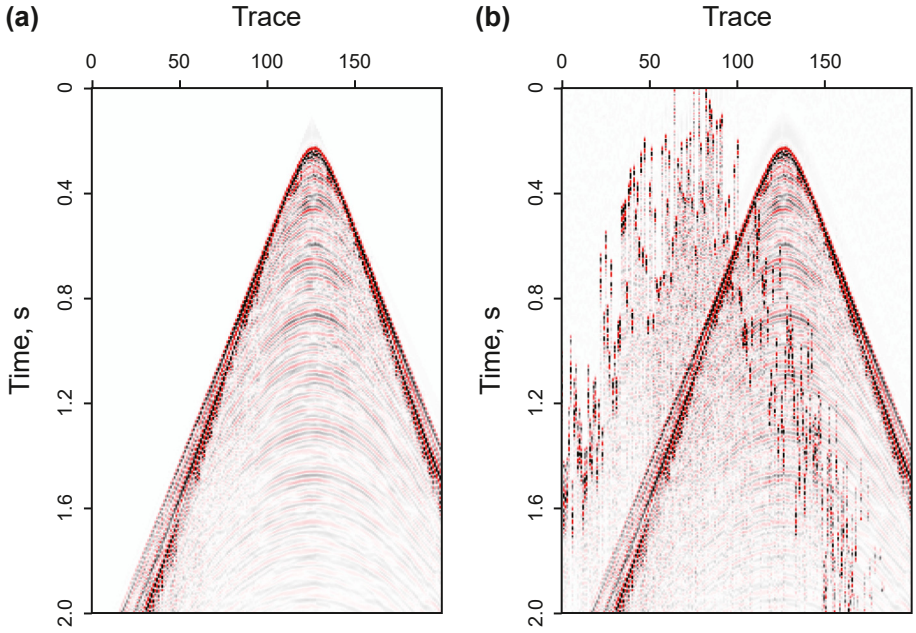

Fig. 6. Field CRG collected from an ocean-bottom-cable survey. **(a)** Unblended CRG. **(b)** Pseudo-deblended CRG.

and the red dashed represents the validation curve.

3.2. Synthetic experiment

In this section, we conduct the comparison experiment among the CT based method, the DT model, the WUDTnet, and the WUDT-NAFnet on complex synthetic CRG. The complex synthetic CRG contains 512 traces and each trace has 1024 sampling points with the time interval of 4 ms. The numerical pseudo-deblended CRG is obtained by a dither time with a range of [−0.4 s, 0.4 s]. To avoid verbosity, only the reference source record is illustrated. As shown in Fig. 3(b), discrete crosstalk interference contaminates the coherent signal. Fig. 4(a)–(d) display the deblended results of the CT based method, the DT model, the WUDTnet model, and the WUDT-NAFnet. We manually remove the blended noise above the first arrival before applying it to each method due to its impact on the deblending results of the CT based method. At first glance, the incoherent interference in Fig. 3(b) can be partially suppressed by these methods. However, upon observing the deblended errors in Fig. 4(e)–(h), we find that the traditional CT based method poorly preserves amplitude at the boundary. The main reason is that the weak coherent signal has the similar curvelet coefficients as the strong incoherent interference. For the DT method, the performance is better than that of the CT based method but still exhibits some weak noise residual. Fig. 4(h) shows the deblended error of the proposed approach, indicating that the error is the minimum among the three deblended errors, highlighting that our hybrid method achieves the best deblending performance. Moreover, the recovered SNRs of the CT based method, the DT model, the WUDTnet, and the WUDT-NAFnet are 16.62, 20.84, 22.46, and 23.51 dB, respectively. It is evident that the designed WUDT-NAFnet excels in handling weak noise and minimizing signal damage.

To intuitively view the damage to the frequency-wavenumber ($f-k$) components of signal, the $f-k$ spectra of the deblended errors of Fig. 4(e)–(h) are estimated. The $f-k$ spectra of the deblended errors of the CT based method, the DT model, the WUDTnet and the presented method are shown in Fig. 5(a)–(d), respectively. It can be observed that the WUDT-NAFnet damages

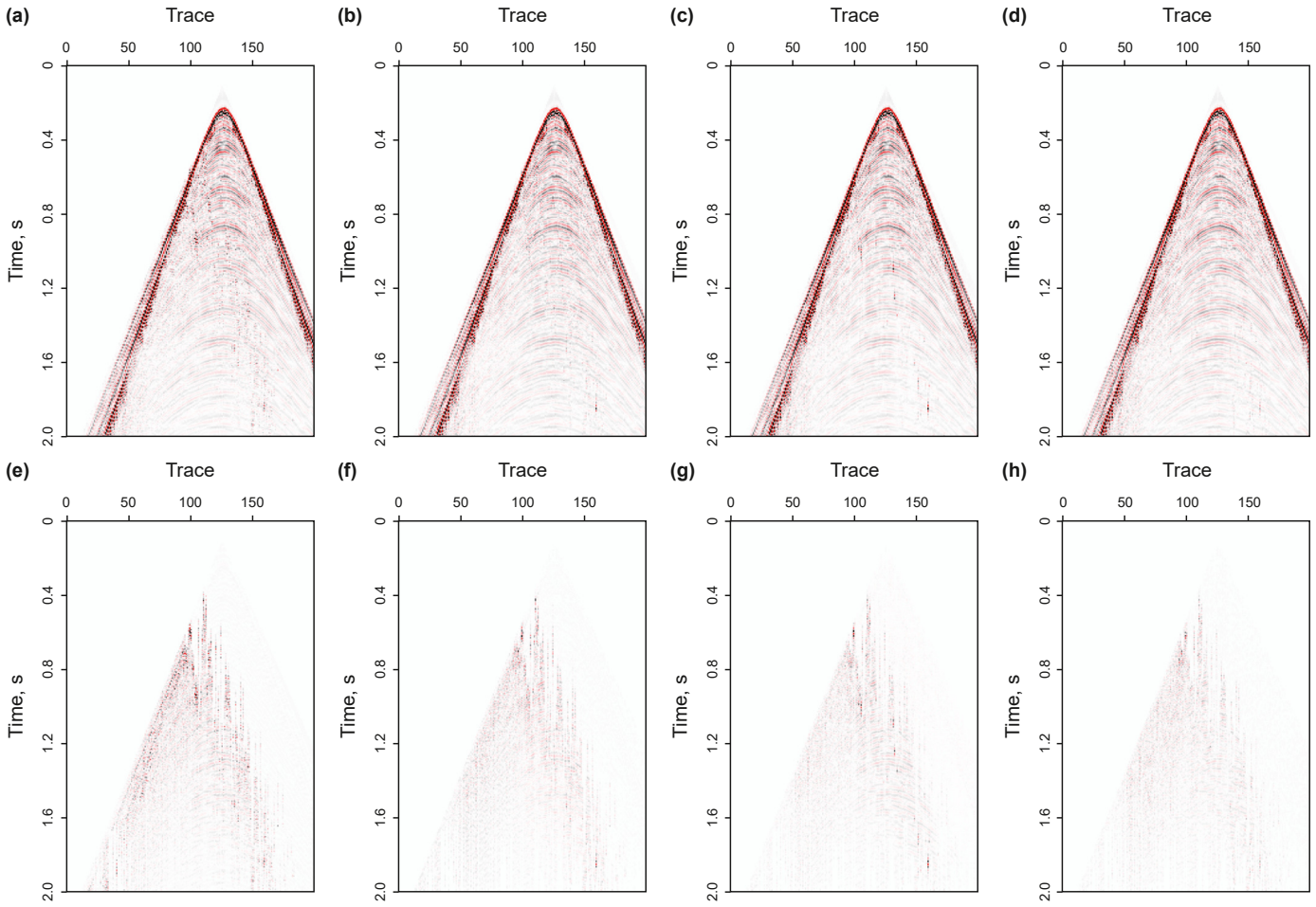

Fig. 7. Deblended results by **(a)** the CT based method, **(b)** the DT method, **(c)** the WUDTnet, and **(d)** the proposed method. Deblended errors related to **(e)** the CT based method, **(f)** the DT method, **(g)** WUDTnet, and **(h)** the proposed method.

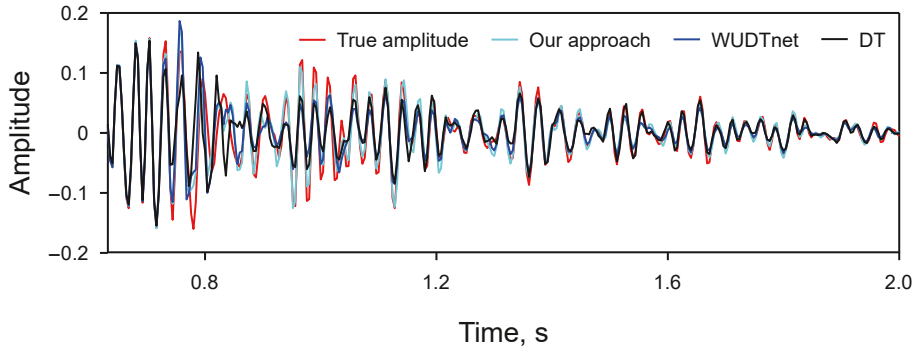

Fig. 8. The 102nd trace record of the first field example, where the red line represents the true amplitude, the cyan line represents the deblended result using our approach, the blue line represents the deblended result using WUDTnet, and the black line represents the deblended result using DT model.

the least amount of signal when compared with the other three methods. This example shows the superiority of our method in retaining useful signal and circumventing amplitude artifacts over the traditional CT based method and the DT method.

3.3. Field data experiments

To further evaluate the effectiveness of the presented algorithm, we compare it with the CT based method, the DT model, and the WUDTnet using two numerical pseudo-deblended field CRG examples. For each field example, we set the dither time to a range of [−0.4 s, 0.4 s]. Moreover, it is worth noting that in our experiments, we didn't employ the strategy of training with additional blended CSGs (Wang and Hu, 2021) or fine-tuning the network parameter using pairs of target zone CRGs (Wang et al., 2021). This omission is intentional to demonstrate the robustness of the proposed WUDTNAFnet.

The first field test CRG is collected from an ocean-bottom-cable

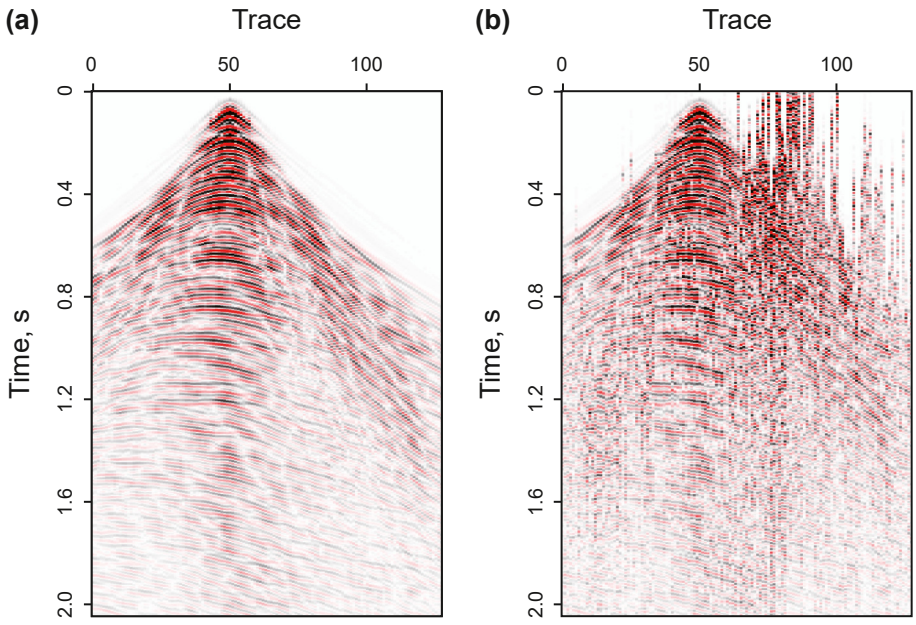

Fig. 9. The 64th CRG of the Gulf of Suez field dataset. **(a)** Unblended CRG. **(b)** Pseudo-deblended CRG.

survey. Fig. 6(a) shows the unblended CRG, and Fig. 6(b) shows the simulated pseudo-deblended CRG of the reference source. From Fig. 5(b), it can be observed that the useful signal of the reference source is contaminated by a large amount of strong crosstalk interference generated from the second source. Fig. 7 illustrates the deblended results and corresponding errors of the CT based method, the DT model, the WUDTnet, and the WUDT-NAFnet, with recovered SNRs of 12.04, 15.95, 17.06, and 18.55 dB, respectively. In addition, the 102nd trace of the true amplitude (the red line), the deblended result by WUDT-NAFnet (the cyan line), the deblended result by WUDTnet (the blue line), and the deblended result by DT (the black line) are shown in Fig. 8. It is obvious that the presented method can restore finer details and more effectively suppress the strong blended noise when compared to the CT based method and the DT model. Furthermore, when compared to the presented WUDTnet, our proposed hybrid method demonstrates superior signal preservation ability. This field example proves that the proposed algorithm can directly apply to separate the field blended CRG and has a more satisfactory deblending performance toward field records than other methods.

Next, we investigate the applicability of the developed approach in field data via experimenting on an open data acquired from the area of Gulf of Suez, which contains 128 gathers and each gather consists of 128 traces with 512 sampling points and the corresponding time interval is 4 ms. To ensure the representativeness of the experimental example, we randomly blended the 64th raw CRG. Fig. 9(b) shows the pseudo-deblended field CRG, because the second source is close to the reference source, the coherent signal is severely obscured by the massive crosstalk noise. The deblended results by the CT based method, the DT model, the WUDTnet, and the proposed approach are shown in Fig. 10(a)–(d), respectively. The corresponding residuals in Fig. 10(e)–(h) show that the proposed approach preserves the most useful components than the other three methods. The recovered SNRs of Fig. 10(a)–(d) are 7.23, 13.43, 15.54, and 16.01 dB, respectively. For the traditional CT based method, the reason for its poor deblending performance is that the coefficient of the curvelet transform is related to the coherency and amplitude value. Although the coherency of the blended noise of the first arrival is poor, its amplitude is large, so the curvelet coefficient difference between the weak coherent signal and the strong incoherent interference is small. In contract, the data-driven based three DL methods can better handle the strong crosstalk noise in this example. Compared to the DT method, the proposed method provides a substantial gain of 2.7 dB. Compared to the WUDTnet, the proposed method has better signal preservation ability, confirming the effectiveness of the WUDT-NAFnet.

To study the generalization ability and flexibility of the presented method, we extract 25 CRGs from the above field dataset with 5-receiver interval for a comprehensive deblending performance analysis. Fig. 11 shows the signal fidelity scores of the different deblending methods, where the black triangle denotes the CT based method, the blue plus stands for the DT method, the green circle represents the WUDTnet model, and the red star represents the proposed method, the proposed method significantly outperforms the other schemes. The average recovery SNRs of the CT based method, the DT model, and the presented approach are 10.38, 14.52, 17.56, and 18.15 dB, respectively. Compared to the CT based method, the presented method achieves 4.28–11.53 dB improvement. For DL based methods, the WUDT-NAFnet yields 2.63–4.45 dB improvement over the DT model, and an improvement of 0.16–1.08 dB over the WUDTnet model. It can be concluded that, the presented algorithm achieves satisfactory deblending performance in the application of field CRGs.

4. Conclusion

In this paper, we present a novel simultaneous source deblending method based on a hybrid WUDT-NAFnet. The WUDT-NAFnet leverages the strengths of both WUDTnet and NAFnet. Specifically, the WUDTnet serves as a powerful filter and is integrated into the iterative estimation and subtraction framework, while the NAFnet is focused on preserving signal amplitude and rejecting interference residuals. Besides, the designed WUDTnet, incorporating discrete Haar wavelet transform and the improved DT block, exhibits favorable performance and faster convergence compared to the DT model. Numerical experiments on synthetic and field CRG examples demonstrate that our method is more effective in deblending than the traditional CT based method and

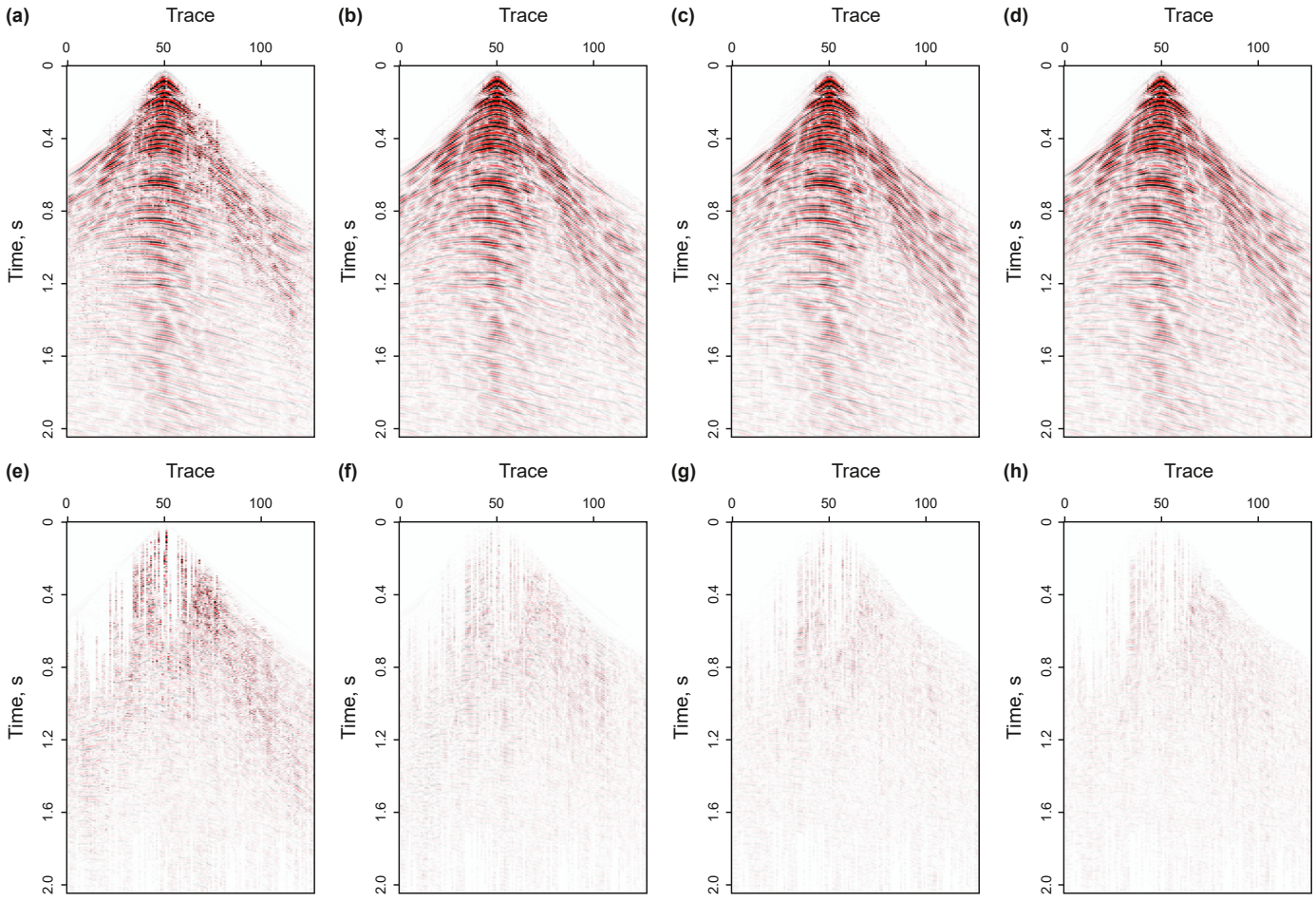

Fig. 10. Deblended results by **(a)** the CT based method, **(b)** the DT method, **(c)** the WUDTnet, and **(d)** the proposed method. Deblended errors corresponding to **(e)** the CT based method, **(f)** the DT method, **(g)** the WUDTnet, and **(h)** the proposed method.

the DT model. Moreover, for field data application, our approach robustly separates the blended interference in numerical field pseudo-deblended CRGs without requiring prior knowledge of the study area's characteristics.

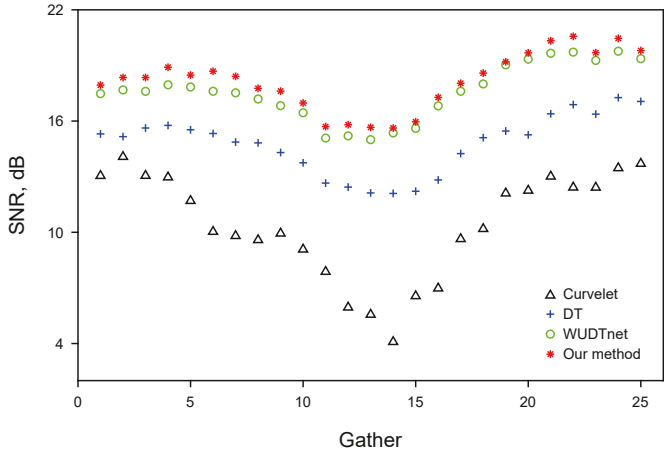

Fig. 11. The recovered SNRs of 25 numerical pseudo-deblended CRGs, where the black triangle is related to the CT based method, the blue plus stands for the DT model, the green circle represents the WUDTnet model, and the red star corresponds to the proposed method.

Declaration of competing interest

The author has no relevant financial or non-financial interests to disclose. There is no conflict of interest to publish this work in this journal.

Acknowledgements

The research is partly supported by the National Natural Science Foundation of China (grant Nos. 42004104, 42030812, 42204136).

References

Chen, L., Chu, X., Zhang, X., et al., 2022. Simple Baselines for Image Restoration. Arxiv Preprint Arxiv:2204.04676.

Chen, Y., Fomel, S., Hu, J., 2014. Iterative deblending of simultaneous-source seismic data using seislet-domain shaping regularization. Geophysics 79, V179–V189. https://doi.org/10.1190/geo2013-0449.1.

Cheng, J., Sacchi, M.D., 2015. Separation and reconstruction of simultaneous source data via iterative rank reduction. Geophysics 80, V57–V66. https://doi.org/10.1190/geo2014-0385.1.

Gan, S., Wang, S., Chen, Y., et al., 2016. Simultaneous-source separation using iterative seislet-frame thresholding. Geosci. Rem. Sens. Lett. IEEE 13, 197–201. https://doi.org/10.1109/lgrs.2015.2505319.

He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. https://doi.org/10.1109/cvpr.2016.90.

Li, C., Huang, J., Li, Z., et al., 2017. Regularized least-squares migration of simultaneous-source seismic data with adaptive singular spectrum analysis. Petrol. Sci. 14, 61–74. https://doi.org/10.1007/s12182-016-0134-1.

Lin, R., Sacchi, M.D., 2020. Separation of simultaneous sources acquired with a high blending factor via coherence pass robust radon operators. Geophysics 85, V269–V282. https://doi.org/10.1190/geo2019-0692.1.

Mahdad, A., Doulgeris, P., Blacquière, G., 2011. Separation of blended data by iterative estimation and subtraction of blending interference noise. Geophysics 76. https://doi.org/10.1190/1.3556597. Q9-Q17.

Nakayama, S., Blacquière, G., 2021. Machine-learning-based data recovery and its contribution to seismic acquisition: simultaneous application of deblending, trace reconstruction, and low-frequency extrapolation. Geophysics 86, P13–P24. https://doi.org/10.1190/geo2020-0303.1.

Qu, S., Zhou, H., Liu, R., et al., 2016. Deblending of simultaneous-source seismic data using fast iterative shrinkage-thresholding algorithm with firm-thresholding. Acta Geophys. 64, 1064–1092. https://doi.org/10.1515/acgeo-2016-0043.

Shi, W., Caballero, J., Huszár, F., et al., 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1874–1883. https://doi.org/10.1109/cvpr.2016.207.

Sun, J., Hou, S., Vinje, V., et al., 2022. Deep learning-based shot-domain seismic deblending. Geophysics 87, V215–V226. https://doi.org/10.1190/geo2020-0865.1.

Sun, J., Slang, S., Elboth, T., et al., 2020. A convolutional neural network approach to deblending seismic data. Geophysics 85, WA13–WA26. https://doi.org/10.1190/geo2019-0173.1.

van Borselen, R., Baardman, R., Martin, T., et al., 2012. An inversion approach to separating sources in marine simultaneous shooting acquisition–application to a gulf of Mexico data set. Geophys. Prospect. 60, 640–647. https://doi.org/10.1111/j.1365-2478.2012.01076.x.

Wang, B., Li, J., Han, D., 2022a. Iterative deblending using MultiResUNet with multilevel blending noise for training and transfer learning. Geophysics 87, V205–V214. https://doi.org/10.1190/geo2021-0341.1.

Wang, B., Li, J., Han, D., et al., 2022b. Deblending and recovery of incomplete blended data via MultiResUNet. Surv. Geophys. 43, 1901–1927. https://doi.org/10.1007/s10712-022-09732-1.

Wang, B., Li, J., Luo, J., et al., 2021. Intelligent deblending of seismic data based on u-net and transfer learning. IEEE Trans. Geosci. Rem. Sens. 59, 8885–8894. https://doi.org/10.1109/tgrs.2020.3048746.

Wang, K., Hu, T., 2021. Deblending of seismic data based on neural network trained in the CSG. IEEE Trans. Geosci. Rem. Sens. 60, 1–12. https://doi.org/10.1109/tgrs.2021.3121806.

Wang, K., Mao, W., Song, H., et al., 2023. A multi-data training method for a deep neural network to improve the separation effect of simultaneous-source data. Geophys. Prospect. 71, 63–84. https://doi.org/10.1111/1365-2478.13276.

Wang, S., Song, P., Tan, J., et al., 2022c. Deblending of seismic data in the wavelet domain via a convolutional neural network based on data augmentation. Geophys. Prospect. https://doi.org/10.1111/1365-2478.13277.

Wapenaar, K., van der Neut, J., Thorbecke, J., 2012. Deblending by direct inversion. Geophysics 77, A9–A12. https://doi.org/10.1190/geo2011-0497.1.

Xu, W., Lipari, V., Bestagini, P., et al., 2022a. Intelligent seismic deblending through deep preconditioner. Geosci. Rem. Sens. Lett. IEEE 19, 1–5. https://doi.org/10.1109/lgrs.2022.3193716.

Xu, W., Zhou, Y., Liu, D., et al., 2022b. Seismic intelligent deblending via plug and play method with blended CSGs trained deep CNN Gaussian denoiser. IEEE Trans. Geosci. Rem. Sens. 60, 1–13. https://doi.org/10.1109/tgrs.2022.3166029.

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.

Xue, Z., Chen, Y., Fomel, S., et al., 2016. Seismic imaging of incomplete data and simultaneous-source data using least-squares reverse time migration with shaping regularization. Geophysics 81, S11–S20. https://doi.org/10.1190/geo2014-0524.1.

Zhang, H., Pang, Y., Liang, S., et al., 2021. Iterative deblending of off-the-grid simultaneous source data. IEEE Access 9, 4923–4938. https://doi.org/10.1109/access.2020.3046638.

Zhang, L., Wang, Y., Zheng, Y., et al., 2015. Deblending using a high-resolution radon transform in a common midpoint domain. J. Geophys. Eng. 12, 167–174. https://doi.org/10.1088/1742-2132/12/2/167.

Zhou, Y., 2017. A POCS method for iterative deblending constrained by a blending mask. J. Appl. Geophys. 138, 245–254. https://doi.org/10.1016/j.jappgeo.2017.01.031.

Zhou, Y., Han, C., Chi, Y., 2018. Deblending of simultaneous-source data using iterative seislet frame thresholding based on a robust slope estimation. J. Appl. Geophys. 153, 17–37. https://doi.org/10.1016/j.jappgeo.2018.02.027.

Zu, S., Cao, J., Qu, S., et al., 2020. Iterative deblending for simultaneous source data using the deep neural network. Geophysics 85, V131–V141. https://doi.org/10.1190/geo2019-0319.1.

Zu, S., Ke, C., Hou, C., et al., 2022. End-to-end deblending of simultaneous source data using transformer. Geosci. Rem. Sens. Lett. IEEE 19, 1–5. https://doi.org/10.1109/lgrs.2022.3174106.