# SEISMIC RANDOM NOISE SUPPRESSION USING DENOISING AUTOENCODER

HUI SONG, MENGHUA FANG, CHENG ZHOU and HOUQIANG GAO

*Sinopec Geophysical Research Institute, Jiangning District, Nanjing 211103, P.R. China. 201400567@yangtzeu.edu.cn*

(Received August 15, 2021; revised version accepted January 4, 2022)

ABSTRACT

Song, H., Fang, M.H., Zhou, C. and Gao, H.Q., 2022. Seismic random noise suppression using denoising autoencoder. *Journal of Seismic Exploration*, 31: 203-218.

In the seismic data acquisition phase, random noise will inevitably be introduced as undesirable components, which is not conducive to subsequent seismic signal processing and imaging tasks. In this article, we propose a denoising framework based on denoising autoencoder (DAE) for seismic random noise suppression. DAE can directly reconstruct noise-free seismic data from noisy seismic data in an unsupervised learning manner. The entire seismic data reconstruction requires three phases: corrupting phase, encoding phase, and decoding phase. In the corrupting phase, the original input is randomly corrupted, which helps the designed network to capture robust features. In the encoding phase, the corrupted input is encoded as a compressed representation that contains the important content of the seismic data. In the decoding stage, the compressed representation is decoded into reconstructed data. We choose the mean square error as the loss function, which is minimized by the back propagation algorithm to update the network parameters. The application of synthetic and real seismic data proves the effectiveness of the proposed method in suppressing seismic random noise.

KEY WORDS: seismic data, random noise, unsupervised learning, denoising autoencoder.

0963-0651/22/$5.00 

## INTRODUCTION

In the process of seismic exploration, seismic signals acquired in the field are usually composed of two kinds of signals: effective signals, and interference signals. Effective signals contain important information about underground structures, which play an important role in oil and gas exploration. Interference signals contain all kinds of interference information unrelated to the strata which makes it more difficult to identify effective signals and brings adverse effects to the interpretation of seismic data. Therefore, suppression of interference signals and improvement of seismic signals quality become an extremely critical step in seismic data processing and interpretation (Lin et al., 2013; Kazemi et al., 2016; Gao et al., 2016; Liu et al., 2018). Interference signals are divided into coherent noise and random noise. Coherent noise has a certain dominant frequency, apparent velocity, and shape, while random noise does not have. Random noise is randomly mixed with the effective signals, which affects the identification of weak signals. Thus, the purpose of this paper is to suppress seismic random noise.

Seismic noise suppression has been an existing problem in geophysics for a long time. Many researchers have put forward a variety of noise suppression methods, which have successfully suppressed seismic noise in practical application. Seismic noise suppression methods can be mainly divided into four categories: prediction-based, sparse transformation-based, decomposition-based, rank-reduction-based. The prediction-based noise suppression methods take advantage of the fact that useful signals are predictable and random noise is unpredictable (Abma and Claerbout, 1995). Sparse transform-based noise suppression methods utilize the characteristics that useful signals and random noises are easily distinguished in the sparse transform domain, such as wavelet transform method (Deighan and Watts, 1997), curvelet transform method (Herrmann et al., 2007; Herrmann and Hennenfent, 2008), and dictionary sparse transform method (Chen et al., 2016a; Siahsar et al., 2017). The decomposition-based noise suppression methods can decompose the seismic data into several different components, and then select the main component representing the effective seismic signals (Chen and Ma, 2014). The principle of rank-reduction-based noise suppression algorithms is that the ideal noise-free seismic data can be constructed into a low-rank matrix, and the appearance of random noise increases the rank of the matrix, so the suppression of random noise is transformed into a matrix rank reduction (Oropeza and Sacchi, 2011; Chen et al., 2016b).

Deep learning is a new branch of machine learning, and its emergence makes machine learning closer to the design goal of artificial intelligence. Its structure consists of multiple processing layers. The features captured at the low-level processing layer are combined into more abstract features at the high-level processing layer. Therefore, deep learning can provide plenty

of abstract features of data layer by layer. In recent years, deep learning has achieved excellent applications in the field of seismic exploration, such as noise suppression (Yu et al., 2019), fault interpretation (Wu et al., 2019), waveform classification and picking (Yuan et al., 2018), and reservoir parameter prediction (Song et al., 2020). There are two strategies for seismic noise suppression based on deep learning: supervised-learning-based and unsupervised-learning-based. The supervised-learning-based denoising method depends on the label data, so the accuracy and quantity of the label data are of great significance to the denoising effect. The unsupervised-learning-based denoising method can denoise the noisy data without label data. Autoencoder (AE) belongs to unsupervised learning, which uses back propagation algorithm to reconstruct unlabeled input. Its goal is to make the output as equal as possible to its input, but it is difficult to extract features that are representative of seismic data. Some improved versions of AE capable of capturing important information of seismic data are proposed, such as Denoising AE (DAE), sparse AE (Zhang et al., 2019; Chen et al., 2019), etc. The DAE is a special neural network with denoising function, which captures features by training corrupted input data. As an unsupervised algorithm, DAE is capable of capturing robust characteristics from unlabeled data.

In this paper, we propose a novel seismic noise suppression framework based on unsupervised learning. Firstly, we introduce the DAE and network parameter optimization. Additionly, we introduce the designed DAE architecture. Finally, we test our proposed method on synthetic seismic data and actual seismic data and compared it with three traditional denoising algorithms. The results show that our proposed method is suitable for seismic random noise suppression.

## METHOD

### Denoising autoencoder

DAE is a special type of deep neural network. DAE is a feature extractor with denoising function, and it can transform a noisy input into a noise-free output. Data reconstruction using DAE requires three stages: (1) corrupting stage; (2) encoding stage; (3) decoding stage. In the corrupting stage, the original input $P$ is first damaged into $\hat{P}$ by the stochastic mapping $q_D$:

$$\hat{P}: \ q_D(\hat{P}|P) \ . \tag{1}$$

In our experiments, we realize the corruption process of input by randomly selecting the components of the input in a fixed proportion to set it to 0, while the others remain unchanged.

In the encoding stage, the damaged input data $\hat{P}$ is encoded into compressed expression $H$:

$$H = \psi_1(w_c \hat{P} + b_c), \tag{2}$$

where $w_c$ and $b_c$ denote weights and bias between the input layer and the hidden layer, respectively; $\psi_1$ is the nonlinear activation function. In our study, the ReLU function $\psi_1(x)$ is used as the nonlinear activation function between the input layer and the hidden layer:

$$\mathrm{ReLU}(x) = \max(0, x). \tag{3}$$

In the decoding stage, the compressed expression $H$ is decoded into reconstructed data $Q$.

$$Q = \psi_2(w_d H + b_d), \tag{4}$$

where $w_d$ and $b_d$ denote weights and bias between the hidden layer and the output layer, respectively; $\psi_2$ is the nonlinear activation function. In our study, the tanh function $\psi_2(x)$ is used as the nonlinear activation function between the hidden layer and the output layer:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{5}$$

Due to the denoising autoencoder is self-supervised learning, the loss function is defined as follows:

$$J(\theta) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{P}\, P - \hat{P}\, \mathrm{P}_2^2, \tag{6}$$

where $\theta$ are the network parameters including weights $\{w_c, w_d\}$ and bias $\{b_c, b_d\}$; $n$ is the sample number. Network parameters $\theta$ play an important role in model performance, and its optimization is achieved by minimizing the loss function.

## Network parameter optimization

In order to better optimize the network parameters, the adaptive moment estimation approach is used in our research (Kingma and Ba, 2015), and its update rules are as follows:

(1) At time step $t$, the gradients $g_t$ should be first computed as

$$g_t = \nabla_\theta J(\theta)_{t\text{-}1}. \tag{7}$$

(2) The biased first moment estimate $u_t$ and the biased second raw moment estimate $v_t$ are computed as

$$u_t = \beta_1 \cdot u_{t-1} + (1-\beta_1)\cdot g_t, \tag{8}$$

$$v_t = \beta_2 \cdot v_{t-1} + (1-\beta_2)\cdot g_t^{\,2}, \tag{9}$$

where $\beta_1$ and $\beta_2$ denote exponential decay rates. Good default parameters are $\beta_1 = 0.9$, $\beta_2 = 0.999$.

(3) The bias-corrected first moment estimate $\hat{u}_t$ and the bias-corrected second raw moment estimate $\hat{v}_t$ are computed as

$$\hat{u}_t = u_t / (1-\beta_1), \tag{10}$$

$$\hat{v}_t = v_t / (1-\beta_2). \tag{11}$$

(4) Finally, parameters $\theta$ are updated as

$$\theta_t = \theta_{t-1} - \eta \cdot \hat{u}_t / (\sqrt{\hat{v}_t} + \varepsilon), \tag{12}$$

where $\eta$ is the learning rate, $\varepsilon$ is a small constant and its default value is $10^{-8}$.

## Proposed DAE architecture

Preparing the training set plays an important role in deep learning. Our training set is obtained from the original seismic data based on the patch-based method. The patch-based method can generate a large number of

training samples, which is conducive to obtaining optimized network parameters. In our study, the patch size is set as 20 × 20. In this paper, we design a denoising autoencoder with a seven-layer structure illustrated in Fig. 1. The seven-layer structure of DAE consists of one input layer, five hidden layers and one output layer. The number of neurons in the input layer and output layer is determined by the patch size, and the number of neurons in the five hidden layers is set to 640, 160, 40, 160 and 640, respectively. It can be observed that DAE is composed of corruption framework, encoder framework, decoder framework. The corruption framework is responsible for randomly corrupting the input. The encoder framework is responsible for encoding the damaged input into a compressed expression. The decoder framework is responsible for reconstructing the compressed expression into an output close to the input.

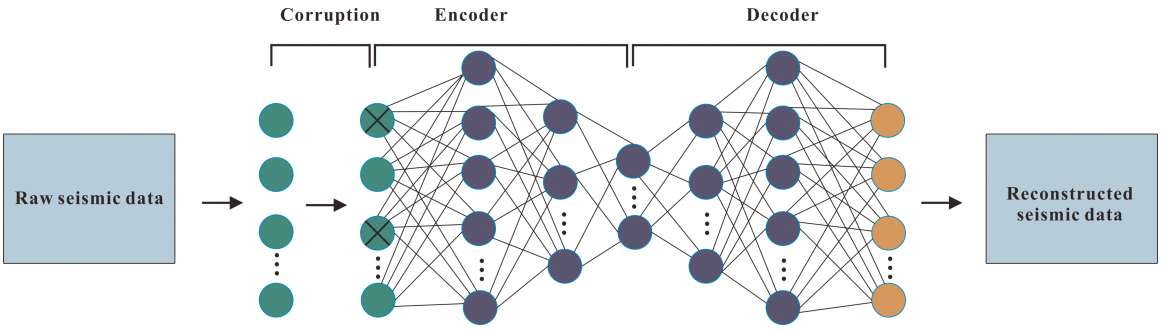


Fig. 1. The proposed DAE architecture.

## EXPERIMENT

We test the denoising performance of the proposed method on two examples. The comparative experiments between the proposed method and traditional methods are done on synthetic and real seismic data. We choose three traditional methods multichannel singular spectrum analysis (MSSA), wavelet transform (WT) and f-x deconvolution (FX), and compare their denoising performance with the proposed method.

### Synthetic seismic data

We first test the denoising performance of four methods on synthetic seismic data. The clean synthetic data consists of 140 traces with 900 time sampling points. Fig. 2a shows the clean seismic data. The synthetic seismic data contains strong and weak events. We add random noise to clean seismic data and obtain noisy seismic data displayed in Fig. 3b. With clean synthetic seismic data available, we can quantitatively evaluate denoising performance based on signal-to-noise ratio (SNR), which is expressed as

$$\mathrm{SNR} = 20\log_{10}\frac{\left\|d_{\mathrm{clean}}\right\|_2}{\left\|d_{\mathrm{denoised}} - d_{\mathrm{clean}}\right\|_2} \quad ,$$

where $d_{\mathrm{clean}}$ and $d_{\mathrm{denoised}}$ represent clean data and denoised data, respectively. Therefore, the SNR of noisy seismic data displayed in Fig. 2b can be calculated to be 1.37 dB.

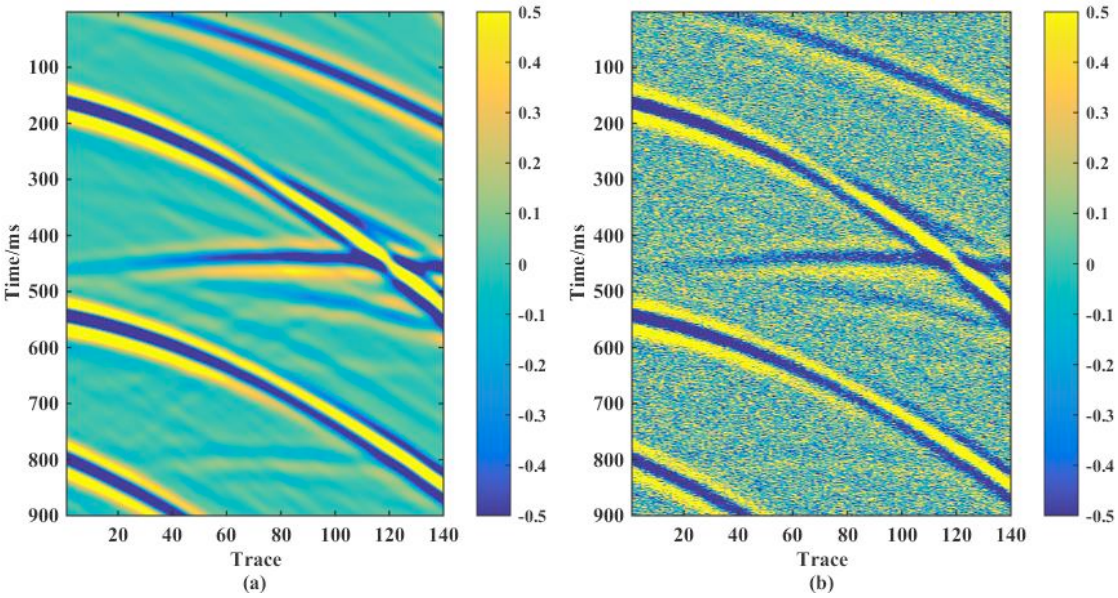


Fig. 2. Synthetic seismic data. (a) Clean Data. (b) Noisy data (SNR = 1.37 dB).

We utilize four denoising methods to denoise the noisy seismic data and obtain the corresponding denoising results illustrated in Fig. 3. It can be observed that the three traditional methods can suppress a lot of random noise, but there is still residual noise in the denoising results. Fig. 3d shows the denoising results of the proposed method , where almost no random noise can be seen. In terms of the SNR of the denoising result, the SNR of the proposed method is 18.34 dB, which is higher than 14.74 dB, 12.86 dB, and 16.77 dB of MSSA, WT, and FX methods, respectively. In order to compare the fidelity of the four methods, we calculated the residual profile of the four methods displayed in Fig. 4. The residual profile is obtained by making the difference between the original seismic data and the denoised results. The ideal residual profile should be chaotic and no trace of useful signal can be observed. It can be seen from Fig. 4 that the useful signal can hardly be seen on the residual profile of the proposed method, which means that the useful signal leakage is the least.

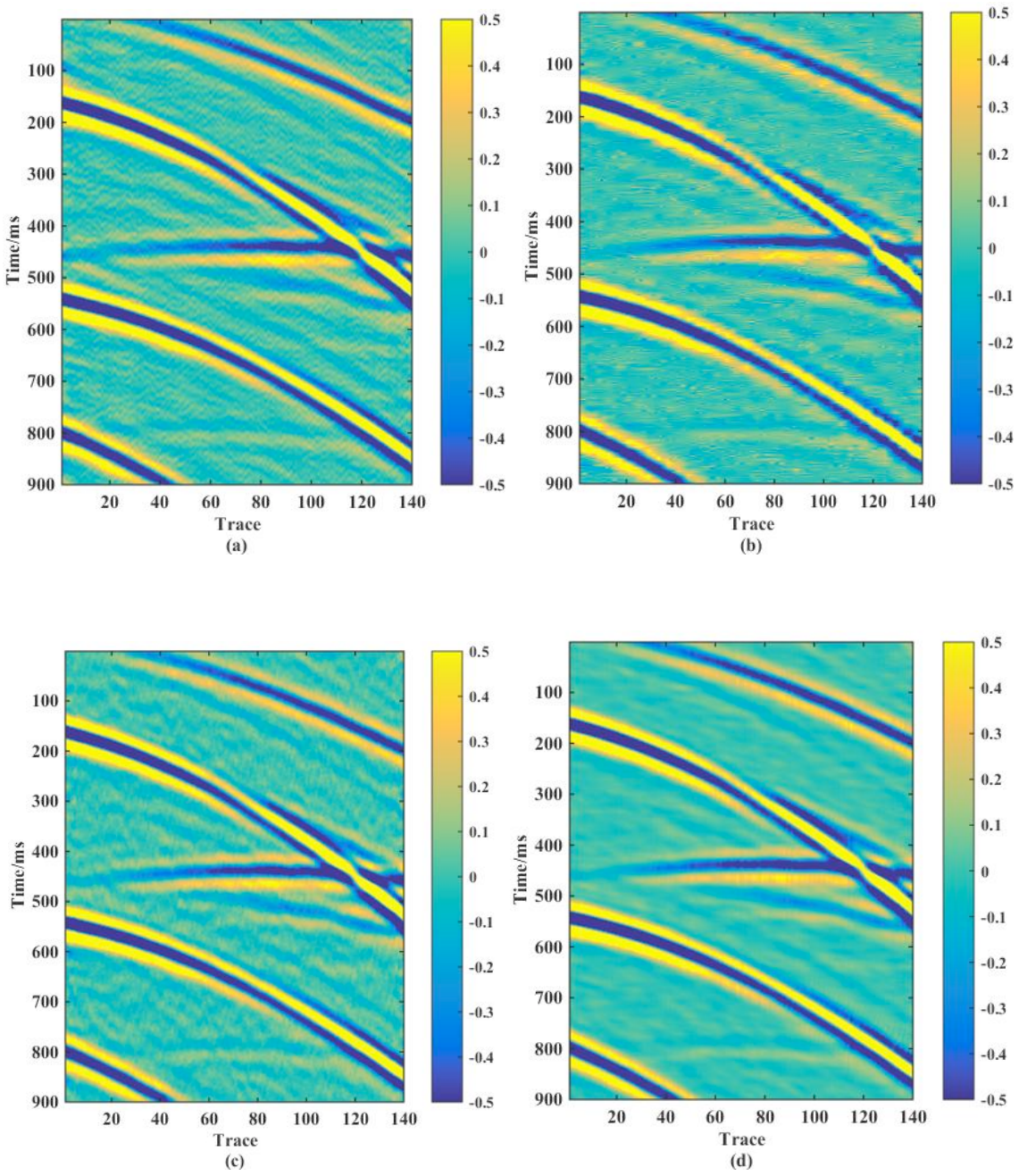


Fig. 3. (a) Denoised results of MSSA (SNR = 14.74 dB). (b) Denoised results of WT(SNR = 12.86 dB). (c) Denoised results of FX (SNR = 16.77 dB). (d) Denoised results of Proposed method (SNR = 18.34 dB).

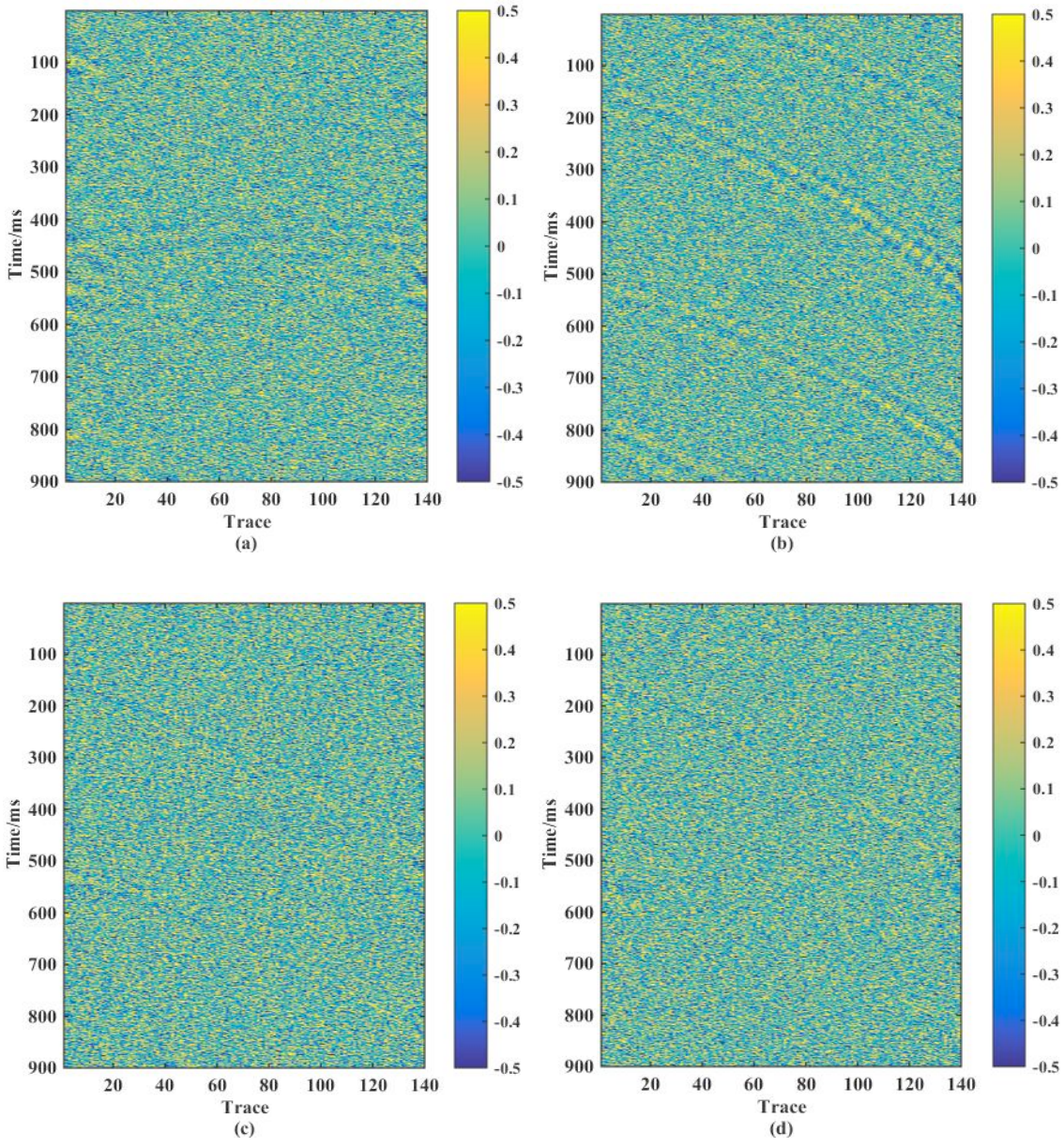


Fig. 4. (a) Residual profile of MSSA. (b) Residual profile of WT. (c) Residual profiles of FX. (d) Residual profiles of Proposed method.

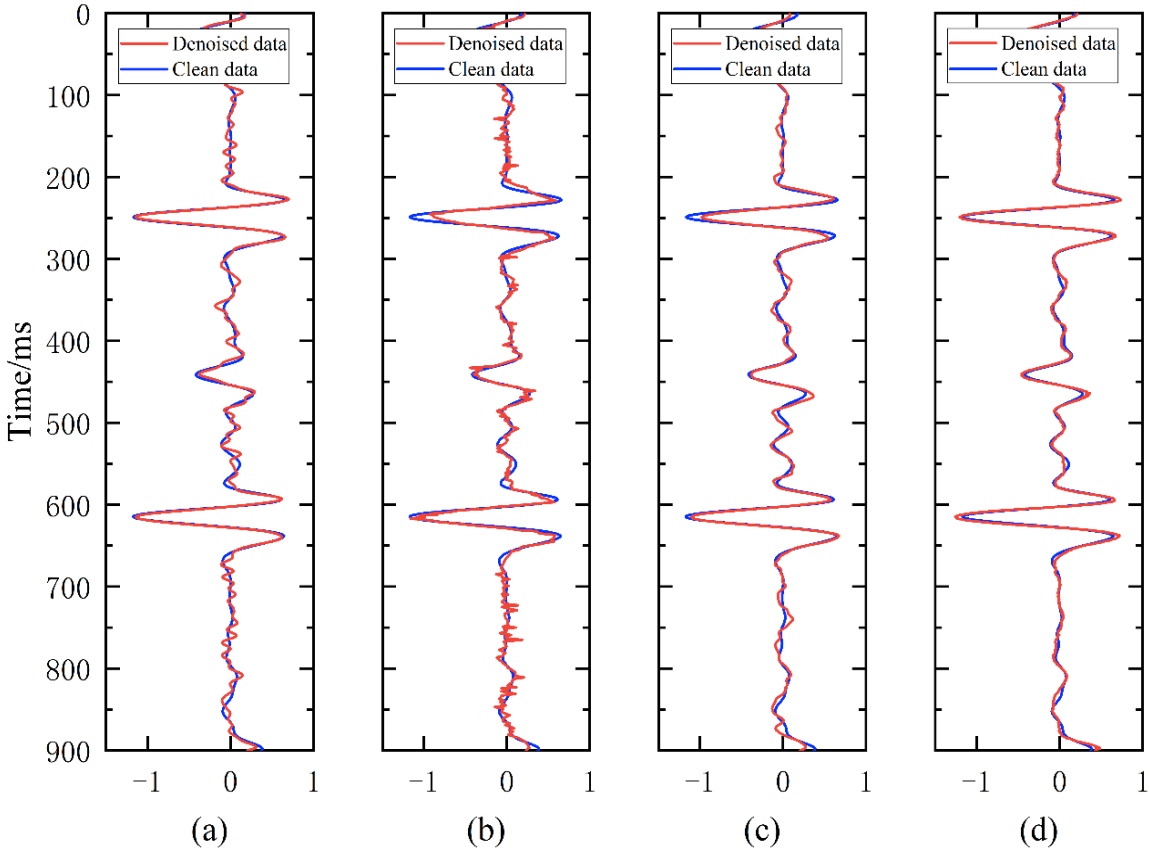


Fig. 5. Denoising results on single channel seismic data. (a) MSSA. (b) WT. (c) FX. (d) Proposed method.

We further compare the denoising results of the four methods on single-trace seismic data, as shown in Fig. 5. It can be observed that the single-trace denoising result of the proposed method is closest to the clean data. In addition, we calculate the SNR of the four methods at different noise levels, as shown in Fig. 6. It can be observed that the SNR of four methods at different noise levels has been improved after denoising, but the proposed method always maintains the highest SNR improvement, which indicates that the proposed method has stronger denoising robustness.

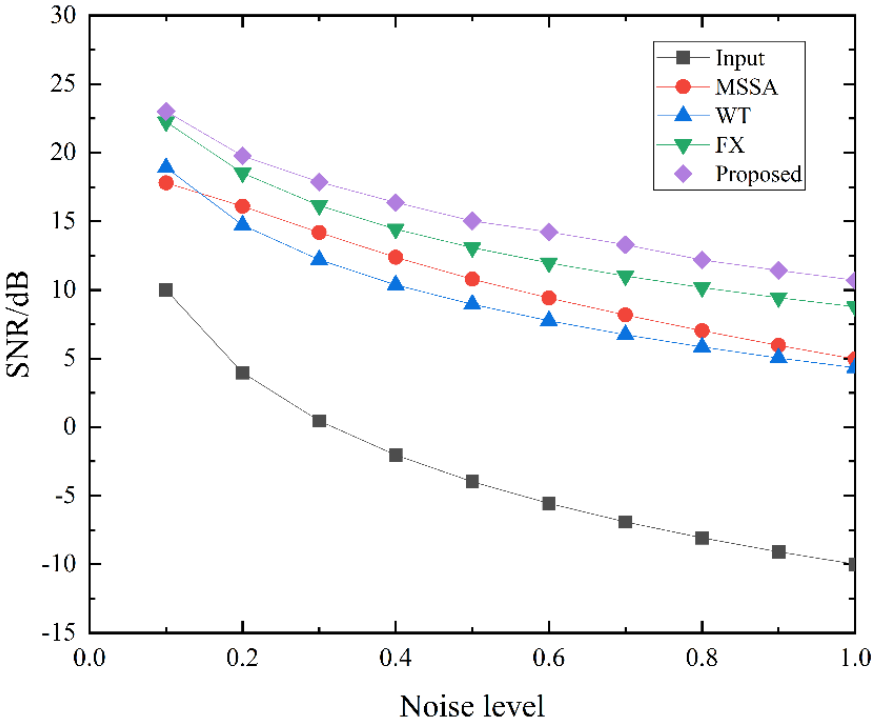


Fig. 6. The SNR comparison of four methods at different noise levels.

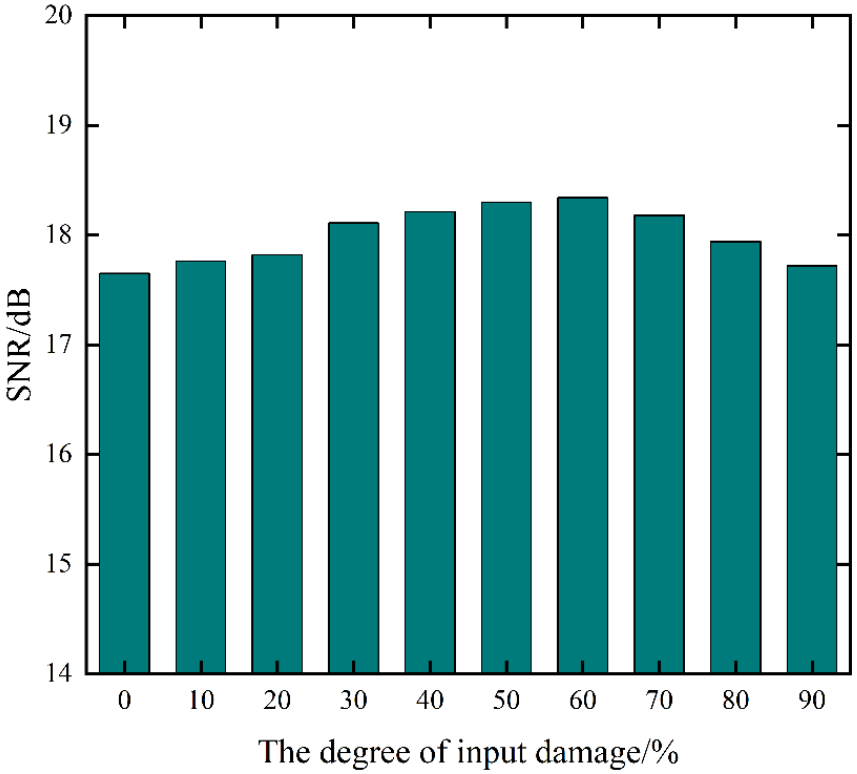


Fig. 7. The effect of input damage on denoising performance.

The effect of input damage on the denoising performance is displayed in Fig. 7. For the proposed method, damage to the input is conducive to the improvement of denoising performance, and the best denoising performance is achieved when the degree of damage is 60%.

## Real seismic data

In the previous section, the proposed method has shown a good denoising effect on the synthetic seismic data. In this section, we assess the denoising ability of the proposed method on the real seismic data. The real seismic data consists of 110 traces with 600 time sampling points. The selected real seismic data are displayed in Fig. 8. It can be seen from Fig. 8 that the real seismic data contains the curve events and faults. Due to the pollution of a large amount of random noise, the continuity of events is poor and weak signals are not clear. The interference of random noise makes the seismic resolution very low.

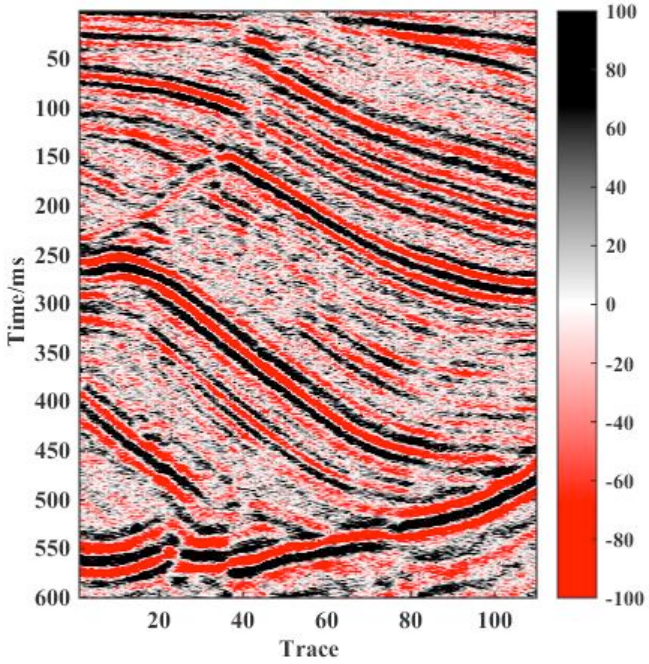


Fig. 8. The real seismic data.

Fig. 9 shows the denoising results of the four methods on the real seismic data. The seismic data denoised by WT has serious distortion of the seismic data, and a lot of local details are lost. There is still a small amount of random noise remaining on the seismic data after denoising by FX. The denoising results of the proposed method are not only more stable in events, but also almost no significant residual noise, which implies a higher profile quality.

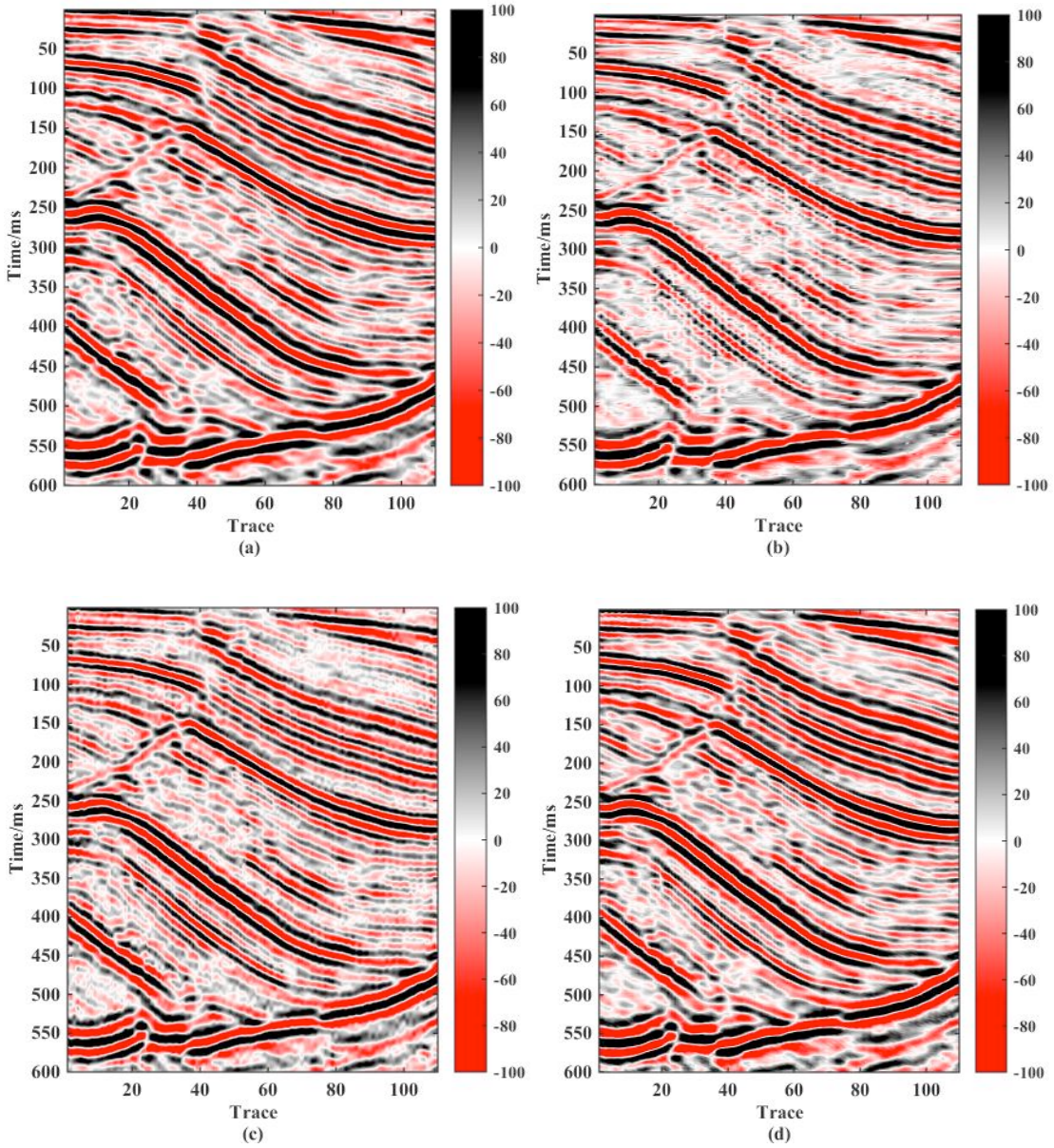


Fig. 9. (a) Denoised results of MSSA. (b) Denoised results of WT. (c) Denoised results of FX. (d) Denoised results of Proposed method.

Since clean real seismic data are unavailable, we can not use SNR as a quantitative standard to compare the denoising performance of the four methods on real seismic data. Thus, we make a rough comparison of the four methods by means of residual profile illustrated in Fig. 10. It is obvious that useful signals can be observed in the residual profiles of the three traditional methods, which means a strong leakage of useful signals. Almost no useful signals can be observed in the residual profile of the proposed method, which means that the proposed method will not cause damage to useful signals.

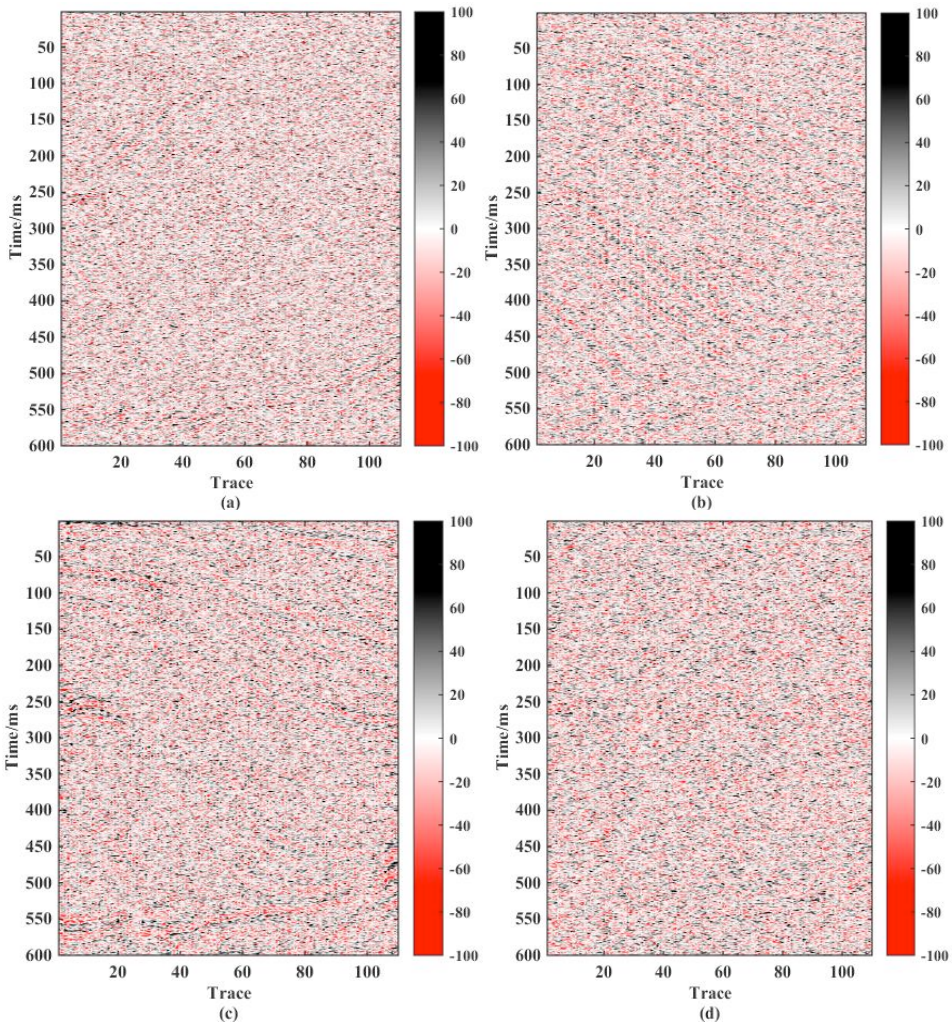


Fig. 10. (a) Residual profile of MSSA. (b) Residual profile of WT. (c) Residual profiles of FX. (d) Residual profiles of Proposed method.

## DISCUSSION

In this section, we compare the calculation time on laptop with Inter(R) Core(TM) i7-6700HQ CPU@2.60GHz and Nvidia GPU GTX960M. On the synthetic data, the calculation time of the proposed method is 12.58 s, while the calculation time of MSSA, WT, and FX method are 0.29 s, 0.64 s, and 0.34 s, respectively, which implies that the proposed method has good denoising performance but takes longer calculation time than the traditional method. In addition, we show the effect of calculation time on denoising performance, as shown in Fig. 11. It can be seen that better denoising performance of the proposed method means longer calculation time. Therefore, how to make a trade-off between denoising performance and calculation time is worthy of further study.

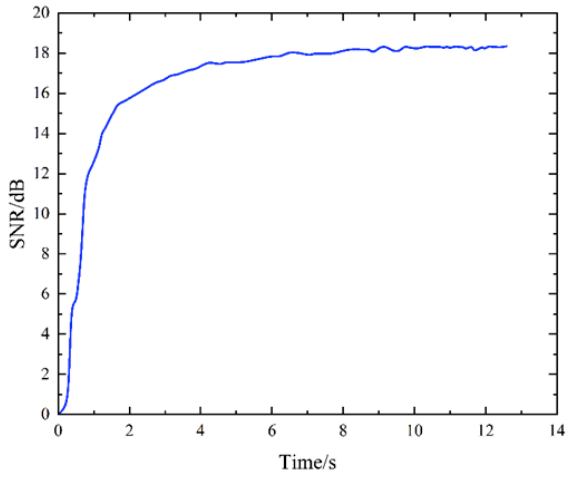


Fig. 11. The effect of calculation time on denoising performance.

The proposed method utilizes the statistical difference between effective signals and random noise to denoise. The proposed method does not have the risk of over-filtering the data and has no special requirements on the data type. However, the proposed method has limitations in removing coherent noise. Since the statistical laws of coherent noise and effective signals are similar, it is difficult for unsupervised-learning-based denoising method to remove coherent noise, which means that the proposed method cannot directly remove coherent noise from unlabeled data. It is feasible to remove coherent noise by introducing label data into the proposed method, which means that the proposed method has changed from unsupervised-learning-based denoising method to supervised-learning-based denoising method.

## CONCLUSION

In order to effectively suppress random noise in seismic data, we have proposed a novel denoising method based on denoising autoencoder. The proposed method can recover clean seismic data from noisy seismic data in an unsupervised manner, which greatly reduces the effort of producing label data. The designed network with corrupting stage, encoding stage, and decoding stage can effectively suppress the random noise in seismic data. Examples with synthetic and real seismic data prove that the proposed method can cause minimal damage to the useful signals while effectively attenuate the random noise of seismic data. Examples also confirm that the proposed method has better denoising performance compared with the three state-of-the-art denoising methods. In the future, we will explore the removal of coherent noise by introducing label data into the proposed method.

# REFERENCES


Abma, R. and Claerbout, J., 1995. Lateral prediction for noise attenuation by t-x and f-x techniques. Geophysics, 60(6): 1887-1896.

Chen, Y. and Ma, J., 2014. Random noise attenuation by f-x empirical mode decomposition predictive filtering. Geophysics, 79(3): V81-V91.

Chen, Y., Ma, J. and Fomel, S., 2016a. Double sparsity dictionary for seismic noise attenuation. Geophysics, 81(2): V103-V116.

Chen, Y., Zhang, D., Jin, Z., Chen, X., Zu, S., Huang, W. and Gan, S., 2016b. Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method. Geophys. J. Internat., 206(3): 1695-1717.

Chen, Y., Zhang, M., Bai, M. and Chen, W., 2019. Improving the signal-to-noise ratio of seismological datasets by unsupervised machine learning. Seismol. Res. Lett., 90(4): 1552-1564.

Deighan, A. and Watts, D., 1997. Ground-roll suppression using the wavelet transform. Geophysics, 62(6): 1896-1903.

Gao, Z., Pan, Z. and Gao, J., 2016. Multimutation differential evolution algorithm and its application to seismic inversion. IEEE Transact. Geosci. Remote Sens., 54(6): 3626-3636.

Herrmann, F., Böniger, U. and Verschuur, D., 2007. Non-linear primary-multiple separation with directional curvelet frames. Geophys. J. Internat., 170(2): 781-799.

Herrmann, F. and Hennenfent, G., 2008. Non-parametric seismic data recovery with curvelet frames. Geophys. J. Internat., 173(1): 233-248.

Kazemi, N., Bongajum, E. and Sacchi, M., 2016. Surface-consistent sparse multichannel blind deconvolution of seismic signals. IEEE Transact. Geosci. Remote Sens., 54(6): 3200-3207.

Kingma, D.P. and Ba, J., 2015. Adam: A method for stochastic optimization. CoRR, abs/1412.6980.

Lin, H., Li, Y., Yang, B. and Ma, H., 2013. Random denoising and signal nonlinearity approach by time-frequency peak filtering using weighted frequency reassignment. Geophysics, 78(6): V229-V237.

Liu, W., Cao, S., Jin, Z., Wang, Z. and Chen, Y., 2018. A novel hydrocarbon detection approach via high-resolution frequency-dependent AVO inversion based on variational mode decomposition. IEEE Transact. Geosci. Remote Sens., 56(4): 2007-2024.

Oropeza, V. and Sacchi, M., 2011. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. Geophysics, 76(3): V25-V32.

Siahsar, M.A.N., Gholtashi, S., Kahoo, A.R., Chen, W. and Chen, Y., 2017. Data-driven multitask sparse dictionary learning for noise attenuation of 3D seismic data. Geophysics, 82(6): V385-V396.

Song, H., Chen, W., Zhang, H., Wang, Y. and Xue, Y., 2020. Sandstone porosity prediction based on gated recurrent units. J. Seismic Explor., 29(4): 371-388.

Wu, X., Liang, L., Shi, Y. and Fomel, S., 2019. Faultseg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. Geophysics, 84(3): IM35-IM45.

Yu, S., Ma, J. and Wang, W., 2019. Deep learning for denoising. Geophysics, 84(6): V333-V350.

Yuan, S., Liu, J., Wang, S., Wang, T. and Shi, P., 2018. Seismic waveform classification and first-break picking using convolution neural networks. IEEE Geosci. Remote Sens. Lett., 15(2): 272-276.

Zhang, M., Liu, Y., Bai, M. and Chen, Y., 2019. Seismic noise attenuation using unsupervised sparse feature learning. IEEE Transact. Geosci. Remote Sens., 57(12): 9709-9723.