Random Noise Attenuation of Sparker Seismic Oceanography Data with Machine Learning

Hyunggu Jun [1], Hyeong-Tae Jou[1], Chung-Ho Kim[1], Sang Hoon Lee[1], Han-Joon Kim[1]

[1]Korea Institute of Ocean Science & Technology, Busan, 49111, Republic of Korea

5   *Correspondence to*: Hyeong-Tae Jou (htjou@kiost.ac.kr)

Abstract. Seismic oceanography (SO) acquires water column reflections by seismic exploration compensating for the drawbacks of conventional physical oceanographic equipment. Most SO studies obtain data using air guns, which have relatively low-frequency bands. For higher-frequency bands at a low exploration cost, using a smaller seismic exploration system, such as a sparker source with a shorter receiver length, would be an alternative. However, the sparker source has a

10   relatively low energy and consequently produces data with a low signal-to-noise (S/N) ratio. To solve the problem of the low S/N ratio of sparker SO data, we applied machine learning. The purpose of this study is to attenuate the random noise in the East Sea sparker SO data without distorting the true shape and amplitude of water column reflections. A denoising convolutional neural network (DnCNN) that successfully suppresses random noise in a natural image is adopted as the machine learning network architecture. One of the most important factors of machine learning is the generation of an appropriate

15   training dataset. We have generated two different training datasets using synthetic and field data. Models trained with the different training datasets are applied to the test data, and the denoised results are quantitatively compared. The trained models are applied to the target seismic data, i.e., the East Sea sparker water column seismic reflection data, and the denoised seismic sections are evaluated. The results show that machine learning can successfully attenuate the random noise of sparker water column seismic reflection data.

20

Ocean Science
Discussions

1 Introduction

Conventional physical oceanography measurements are performed by dropping equipment at the observation points. In general, due to time and cost limitations, the distance between observation points is large, from hundreds of meters to tens of kilometers; thus, the acquired water column information has a low horizontal resolution. Holbrook et al. (2003) suggested a seismic oceanography (SO) method that obtained water column reflections via seismic exploration and analyzed seismic sections to estimate the oceanographic characteristics of sea water, and they successfully imaged the Atlantic oceanographic structure. Seismic exploration acquires data continuously in the horizontal direction; thus, it has the advantage of generating data with a high horizontal resolution compared to conventional oceanographic methods. Therefore, SO is widely used to identify the structure of water layers (Tsuji et al., 2005; Sheen et al., 2012; Piété et al., 2013; Moon et al., 2017), estimate physical properties (i.e., temperature, salinity) (Papenberg et al., 2010; Blacic et al., 2016; Dagnino et al. 2016; Jun et al., 2019) and determine the behavior of turbulence or internal waves (Sheen et al., 2009; Holbrook et al., 2013; Fortin et al. 2016).

 SO has been conducted mainly using air guns, a high-energy source, and the central frequency of air guns is usually below 100 Hz, which is relatively low. Therefore, the vertical resolution of the acquired seismic data using air guns is several meters, which is lower than that of conventional physical oceanography observation equipment. SO also has the disadvantage of higher exploration expenses when using air guns and streamers that are several kilometers long. Ruddick (2018) highlighted the limitations of current SO studies using multichannel seismic (MCS) exploration and argued that using a small-scale source instead of a large-scale air gun and a relatively shorter streamer with a length shorter than 500 m can make SO more widely available.

 Piété et al. (2013) implemented a sparker source with a central frequency of 250 Hz and a short 450-m streamer (72 channels at 6.25 m intervals) to examine the oceanographic structure. Since relatively high-frequency band sources were implemented, data with a high vertical resolution of 1.5 m were acquired, and the short source signature enabled the thermocline structure to be imaged even in very shallow areas between 10 and 40 m. However, the signal-to-noise (S/N) ratio of the seismic section was lower than that of the air gun source, and the amplitude of the thermocline feature was small; thus, it was difficult to interpret. Generally, using a low-energy source and a short streamer in seismic exploration causes the low-S/N ratio problem. This problem becomes more serious in seismic exploration targeting the water layer because the difference in impedance between layers is smaller than that with the subsurface. If a low-energy source is used, the water column reflections recorded by the receiver become too weak, and the background noise becomes relatively large. The improvement in vertical resolution is evident when using higher-frequency band sources such as a sparker source; therefore, if appropriate methods can effectively suppress the random noise in the seismic section, more useful information can be derived compared to SO data using an air gun source.

 There are various types of noise recorded by the receiver in seismic exploration, and several data processing steps are usually applied to the seismic data to attenuate noise. However, the noise attenuation method not only removes noise but also potentially alters important seismic signals (Jun et al., 2014). Especially for SO data, careful processing is essential to recover

55   the actual shape of the water column reflections (Fortin et al., 2016), which contain internal wave and turbulence information. It is difficult to apply various noise attenuation methods to SO data because analyzing the internal wave and turbulent subranges of the water column requires the horizontal wavenumber spectrum (Klymak and Moum, 2007) of the seismic data, which is liable to be damaged by data processing. Therefore, minimized noise attenuation processes have been applied to SO data, and for this reason, studies calculating the wavenumber spectrum by using SO data such as those by Holbrook et al. (2013) and

60   Fortin et al. (2016, 2017) have only applied bandpass and notch filters to remove random and harmonic noise. However, when the sparker is used as a seismic source, the bandpass filter alone is not sufficient to attenuate random noise, resulting in great difficulties in analyzing the wavenumber spectrum. Therefore, it is necessary to apply additional data processing to properly attenuate noise without damaging the wavenumber characteristics of SO data.

Recently, rapid advances in computer resources have spurred artificial intelligence (AI) research, and several studies have

65   been conducted to apply machine learning in the field of seismic data processing (Araya-Polo et al., 2019; Yang and Ma, 2019; Zhao et al., 2019). Among them, one of the most actively studied areas is prestack and poststack data noise attenuation. Noise attenuation using machine learning has been widely studied, such as the multilayer perceptron (MLP) (Burger et al., 2012) and autoencoder (AE) (Xie et al., 2012; Wu et al., 2016). After convolutional neural networks (CNNs) were introduced, various noise attenuation methods based on the CNN architecture have been proposed (Jian and Seung, 2009; Gordonara, 2016;

70   Lefkimmiatis, 2017), and the denoising convolutional neural network (DnCNN) suggested by Zhang et al. (2017) attained good results in random noise suppression in natural images. The DnCNN was applied to attenuate various types of noise from seismic data such as ground roll (Li et al., 2018) from onshore field prestack seismic data and random noise from synthetic prestack seismic data (Si and Yuan, 2018) and three-dimensional field seismic cubes (Liu et al., 2018). The DnCNN uses residual learning (He et al., 2016) and has the advantage of minimizing damage to the seismic signal by estimating the noise

75   from seismic data rather than directly analyzing the signal. The original shape of the water column reflector in SO data remains unchanged during data processing, so the DnCNN, which learns noise characteristics, is a suitable SO data denoising algorithm. Therefore, this study applies the DnCNN to attenuate random noise in East Sea sparker SO data.

As important as the proper neural network architecture when conducting training through machine learning is the use of an appropriate training dataset. When using the DnCNN to attenuate noise, the training data require noise-free and noise-only (or

80   noise containing) data. In this study, we use both field and synthetic data as training data and compare which training data are more suitable for the DnCNN in attenuating random noise in SO data.

First, we introduce the DnCNN architecture used in this study and explain the construction method for the training and test datasets using field and synthetic data, respectively. Then, we perform training using the constructed training datasets and verify the trained models using test datasets. Finally, the trained models are applied to the East Sea sparker SO data, and the

85   results are compared and evaluated.

Ocean Science
Discussions
Open Access

EGU

2 Data and Methodology

2.1 Review of the DnCNN

The purpose of this study is to attenuate the random noise in sparker SO data, and the machine learning architecture used in
90 this study is the DnCNN, which was suggested by Zhang et al. (2017). DnCNN is a neural network architecture based on the
CNN for the purpose of removing the random noise in natural images. DnCNN reads the noisy image in the input layer and
extracts the noise from the noisy image during the hidden layer. Layer is a module containing several computing processes
(e.g. convolution, pooling or activation). At the output layer, the extracted noise is subtracted from the noisy image and
generates the denoised result. The architecture of the DnCNN is shown in Fig. 1 and will be explained in more detail below.
95 The DnCNN has three distinctive characteristics: 1) residual learning, 2) batch normalization, and 3) the same input and output
data size for each layer.

Residual learning was first suggested by He et al. (2016) and it added the shortcut connection to the neural network to
overcome the problem of machine learning when networks delve deeper. The DnCNN adopted residual learning and a single
shortcut to estimate the noise from natural images. The estimated noise was subtracted from the noisy natural image, and the
100 noise-attenuated image remained. If the DnCNN is applied to seismic data denoising, the target noise is estimated from the
noisy prestack or poststack seismic data, and the estimated noise is subtracted from the noisy seismic data. The seismic data
including noise (y) can be expressed by adding noise-free seismic data (x) and noise (n) as follows:

$$y = x + n. \tag{1}$$

When the deep learning architecture that estimates noise from the noisy seismic data is $\boldsymbol{D}(y; n)$, the cost function of the
105 DnCNN (C) can be expressed as follows:

$$C = \frac{1}{2N} \sum_{i=1}^{N} \|\boldsymbol{D}(y_i; n_i) - (y_i - x_i)\|^2, \tag{2}$$

where n is the estimated noise from the original noisy seismic data (y), N is the number of the training data and $\| \quad \|^2$ is the
sum of squared errors (SSE). Although the DnCNN uses residual learning, it is different from the conventional residual network.
The conventional residual network utilizes residual learning to solve the performance degradation problem when the network
110 depth increases; thus, it includes many residual units. On the other hand, the DnCNN uses residual learning to predict noise
from noisy images, which is related to trainable nonlinear reaction diffusion (TNRD) (Chen and Pock, 2016) and includes a
single residual unit. For example, ResNet (He et al., 2016), which is a well-known image recognition network using residual
learning, has more than tens or hundreds of network depth layers, but the DnCNN has fewer than 20 network depth layers.
Moreover, the DnCNN applies batch normalization (Ioffe and Szegedy, 2015) after each convolution layer to transform the
115 mini-batch data distribution. The distribution of input data varies during training and the neural network has a risk of updating

the weights to the wrong direction. Batch normalization is a method to normalize the distribution of each mini-batch by making the mean and variance of the mini-batch as 0 and 1, respectively. The normalized mini-batch is transformed through scaling and shifting. The. Batch normalization is widely used in many deep learning neural networks because it can stabilize learning and increase the learning speed (Ioffe and Szegedy, 2015). The authors of the DnCNN empirically found that residual learning

120   and batch normalization create a synergistic effect. In addition, unlike the encoder-decoder type denoising architecture, the size of the input data of the DnCNN is the same as the size of the output data in each layer. The DnCNN directly pads zeros at the boundaries during convolution and does not contain any pooling layer; thus, the data size remains unchanged during training. This procedure has the advantage of minimizing the data loss occurring during the encoding and decoding process. The amplitude and shape of the seismic reflections are important for spectrum analysis using SO data. To minimize possible

125   deformation of the seismic signals during the denoising procedure, the DnCNN, which predicts noise using residual learning and avoids information loss due to the absence of an encoding-decoding model, could be an appropriate algorithm.

2.2 Network architecture

The DnCNN uses three different kinds of layers, and we use the same layers as suggested by Zhang et al. (2017). The first

130   layer type consists of "convolution + rectified linear units (ReLUs; Krizhevsky et al. (2012))" and is used only at the first layer. In the convolution process, 2-dimensional convolution between the certain size of kernel and data is performed. The outputs of the convolution process are pass through the activation function to add the nonlinearity in the network. ReLU is used for the activation function in this study. The size of the convolution filter is $3 \times 3 \times c$ and generates 64 feature maps, where c is the number of channels of the input data. The conventional DnCNN performs denoising from the image file (.jpg, .png, etc.),

135   and the size of the convolution filter is $3 \times 3 \times 3$ in the color image and $3 \times 3 \times 1$ in the gray image. On the other hand, this study extracts noise from binary files, and a $3 \times 3 \times 1$ convolution filter is adopted. The second layer type consists of convolution + batch normalization + ReLUs and is applied from layers 2 to L-1, where L is the total number of network layers. Sixty-four $3 \times 3 \times 64$ convolution filters are used because the number of feature maps of hidden layer is 64 which is the same at all hidden layers. After convolution, batch normalization and the ReLU activation function are applied. The third layer type

140   is convolution and uses only the last layer to generate output noise data, and one $3 \times 3 \times 64$ convolution filter is used. After training is completed, the predicted noise is subtracted from the input data to produce denoised data. Fig. 1 shows the DnCNN architecture used in this study, where Conv and BN indicate convolution and batch normalization, respectively.

2.3 East Sea SO data

145   The purpose of this study is to attenuate the random noise in the East Sea sparker SO data. The East Sea sparker SO data were obtained with a 5,000-J SIG PULSE L5 sparker source to investigate the propagation of the internal tide and characteristics of turbulent mixing. Two seismic lines were explored: line 1 traveled from southwest to northeast, and line 2 traveled from

northeast to southwest (Fig. 2). The seismic data include the shallow continental shelf and slope with a water depth of ~ 200 m, but we removed the continental shelf and slope area and used 280.4 km of line 1 and 280.9 km of line 2 because the data

150 from these sections did not target the subsurface but the water layer. The shot interval was approximately 15 m, and 24 receivers were used at intervals of 6.25 m.

 The acquired seismic data were processed through conventional time processing consisting of instrument delay and spherical divergence corrections, bandpass filter, common-midpoint (CMP) sorting and stacking. Spherical divergence correction was performed by empirically multiplying the square root of time at each time step. The corner frequencies of the trapezoidal

155 bandpass filter (Dickinson et al. 2017) was 60-80-250-300 Hz, which was higher than that in air gun seismic data processing. Sparker source data have a lower S/N ratio due to the weak energy source compared to air gun source data and generally rely on a shorter streamer length; thus, it is common to generate supergathers (Piété et al., 2013) to enhance the S/N ratio. We combined 4 neighboring CMP gathers (Tang et al., 2016) to construct one supergather. A constant velocity of 1,500 m/s was adopted for normal move-out. After CMP stacking, data recorded before 0.03 s were eliminated from the stack section because

160 only direct waves and noise were present, and water layer reflections were rarely recorded. The processed seismic sections are shown in Fig. 3. The internal wave of the research area propagates above a depth of 200 m, which is approximately 0.26 s in the seismic section. In addition, the physical properties of the research area were measured with oceanographic equipment, such as conductivity/temperature/depth (CTD), expendable conductivity/temperature/depth (XCTD) and expendable bathythermograph (XBT) instruments, during exploration. Fig. 4 shows the temperature profiles from two XBT casting

165 locations. From the measurement data, the mixed layer ranged from the sea surface to a depth of 30 m, the depth of the thermocline ranged from 30 to 200 m and deep water occur below approximately 200 m depth. Fig. 5 shows the reflection coefficients calculated with the XBT data and assuming a constant density. The reflection coefficients are very small at depths shallower than 30 m, which seems to be the mixed layer, and deeper than approximately 200 m, which seems to be the deep water layer. Deep water exhibits a very slight water temperature/salinity variation with the depth, which makes it difficult to

170 generate reflections, as indicated by the seismic sections and reflection coefficients. Therefore, data after 0.28 s are considered random noise, and we used this part as noise data for the DnCNN.

2.4 Training data

The most important noise attenuation aspect of machine learning is generating an appropriate training dataset. Noise-free

175 seismic sections (the ground truth) and sections with noise are required to generate the training dataset, and the training dataset can be constructed by combining these two datasets. As previously explained, the purpose of this study is to effectively attenuate noise in the water column seismic section acquired in the East Sea. Thus, the noisy section can be easily obtained by extracting the deep water zone of the water column seismic section without reflections. At this point, we assume that the random noise of the top and bottom parts of the water column seismic section exhibits similar features. The sections with noise

180 extracted from the East Sea SO data are shown in Fig. 6. There are no notable reflections in the sections with noise. However,

it is almost impossible to obtain noise-free seismic sections from field data. Therefore, we constructed training datasets using two different methods and compared these datasets.

Training dataset 1 obtains the ground truth based on the field subsurface sparker seismic section. The reflection coefficient of the subsurface is tens to hundreds times larger than that of the water column; thus, the subsurface seismic data have a better

185 S/N ratio than the SO data. In addition, after the proper data processing steps, the S/N ratio of the subsurface seismic data can be further enhanced. We used 14 lines of subsurface field sparker seismic data (SEZ data) acquired with the same equipment used to record the East Sea SO data. We used the interval from 0.2 to 0.6 s of the original data where the noise level is relatively low. A bandpass filter, FX-deconvolution, a Gaussian filter and noise muting above the sea floor were applied. Fig. 7 (a) shows an example of the sparker subsurface seismic data used to generate training dataset 1. This method has the advantage of using

190 data with similar characteristics to those of the target data (the East Sea SO data) as the ground truth because the data are field data recorded with the same equipment. Even if the S/N ratio of the sparker subsurface seismic data is relatively higher than that of the sparker SO data and noise is suppressed during processing, it is difficult to completely eliminate noise from seismic data. Therefore, this method has the disadvantage that there is a possibility that the remaining noise would have a detrimental effect on training.

195 Training dataset 2 uses synthetic data as the ground truth. The method for generating a synthetic seismic section from the velocity model is to perform time or depth domain processing using prestack synthetic data or to convolve the reflection coefficient and source wavelet. The former method has the advantage of generating synthetic seismic sections with features more similar to those of the actual field seismic section, but the generation and processing of prestack data are time consuming, and artificial noise is often generated during processing. The latter method has the advantage of generating noise-free seismic

200 sections with a very simple procedure. However, the generated synthetic seismic section has much different features from the target seismic section, which is, in this study, the East Sea water column sparker seismic section. Therefore, when the trained model is applied to the target seismic section, there is a risk that the trained model will regard the reflection signal as noise. In this study, we used the latter method to generate the ground truth because we needed to avoid the artificial noise. Marmousi-2 and Sigsbee2a synthetic velocity models with a constant density (1 g/cc) were employed to calculate the reflection coefficient,

205 and the first derivative Gaussian wavelet was the synthetic source wavelet. The original Marmousi-2 and Sigsbee 2A synthetic velocity models are depth domain velocity models, but we assumed that these velocity models were time domain models to generate time domain seismic sections via 1-dimensional convolutional modeling. Fig. 7 (b) and (c) show the generated seismic sections of Marmousi-2 and Sigsbee 2A, respectively.

Each ground truth was first divided into with 300×300 sections. Then, amplitude values higher than the top 1% and lower

210 than the bottom 1% were replaced by the top 1% and bottom 1% values, respectively, to prevent outliers from significantly affecting training. In addition, the outlier-removed ground truth and noise section were normalized to the maximum value of each section. This procedure balances the amplitudes of the ground truth and noise before generating the training dataset. Finally, training data with field seismic noise were generated by combining the ground truth and noise at a random ratio. Eq. 3 is the method to construct the training data, and Fig. 8 shows an example training data compilation.

$$T = r_1 \times G + r_2 \times N,$$

(3)

215

where T is the noise-added seismic patch (training data), G is the ground truth patch, N is the noise patch extracted from the noisy part of the East Sea seismic section (noisy data), and r_1 and r_2 are random values ranging from 0.3~0.8 ($r_1 + r_2 = 1$). The dimensions of T, G and N are 50×50; G and N were extracted at a random location of the ground truth and noisy section. To increase the ground truth data size, data augmentation was applied by zooming in/out and randomly rotating or flipping the

220   data. Training data were newly generated at every epoch with the fit_generator function in Keras (Keras Documentation, 2020). The fit_generator function generated 28,160 data points at every epoch because the mini-batch size was 128 and the iteration of each epoch was 220. The epoch means the number of processes using entire data point and iteration means the number of processes using a mini-batch in a epoch.

Ocean Science
Discussions

3 Training

3.1 Experimental setting

The experiment was conducted using 28,160 training data points per epoch, and the size of each patch was 50×50. The mini-batch size was 128, the network depth which is the total number or layers in the network architecture was 17, the number of feature maps of each layer was 64 and the Adam optimizer (Kingma and Ba, 2015) was implemented by following Zhang et al. (2017)'s DnCNN experiments. The network architecture used in this study is shown in Fig. 1. We performed training by using the two different training datasets generated from the field data (training dataset 1) and synthetic data (training dataset 2). The DnCNN model was trained for 40 epochs, and the total training time was approximately 1 hour using a single NVIDIA Quadro P4000 GPU.

3.2 Experiment using training dataset 1

Training dataset 1 was generated with the SEZ field data and noise obtained from the East Sea seismic section. After training the DnCNN model (D1 model) using training dataset 1, we evaluated the trained model against the test data. The test data were generated with the same procedure as that for the training data, and we used the other lines of SEZ data which were not used to generate the training data. The number of 300×300 size test data was 86 and we divided the test data into (50×50) size patches which is the same size as the training data patch. Then, we discarded the remaining data divided by 128 (mini-batch size) for the computational efficiency, thus the number of test data points was 3,072. Fig. 9 shows 6 randomly selected test data, ground truth and denoised results after applying the D1 model at the 5th, 10th, 20th and 40th epoch. The depicted test data (1 to 6) include noise, but most of the noise has been successfully removed after training for 40 epochs. Especially in test data 3 and 6, the reflections are hardly recognized because of the severe noise, but the D1 model successfully attenuated the noise and generated a denoised section almost identical to the ground truth. In addition, there is a water layer without any signal at the top of test data 4, and the trained model properly attenuated the noise at the water layer. This means that the trained model can determine those parts where no signal occurs.

However, the trained model using training dataset 1 has one problem. The ground truth of test data 5 contains noise in the right bottom part, and training dataset 1 also contains noise in some parts of the ground truth. Even though the ground truth of training dataset 1 was generated from a processed subsurface sparker seismic section, which had a relatively high S/N ratio, noise remained because it is almost impossible to perfectly remove noise from field data. The ground truth of training dataset 1 (SEZ data in Fig. 7 (a)) is obtained using the same equipment as was used for the East Sea SO data, which is the target of this study. Therefore, the ground truth signal has similar characteristics to the signal of the East Sea SO data, but its noise feature could also be similar to the noise of the East Sea SO data. This means that noise with similar characteristics would be trained to be eliminated in some cases and not in other cases during training. Training inconsistency can degrade the performance of the trained model.

To evaluate the test result quantitatively, we calculated the peak S/N ratio (PSNR) and structural similarity index measure (SSIM). The PSNR reflects the amount of noise contained in the data and can be calculated as follows (Hore and Ziou, 2010):

$$PSNR = 20 \, log_{10}(MAX_I) - 10 \, log_{10}(MSE) \qquad (4)$$

260

where MAX_I is the maximum value of the image and MSE is the mean squared error between the data with and without noise. The PSNR is high when noise is successfully removed, while the PSNR is low when noise is not sufficiently removed. Fig. 10 (a) shows the average PSNR of the test results. At the early stage of training, the average PSNR is low, which indicates that noise has not been sufficiently removed, but it increases as training progresses and converges at approximately 36 dB after 25

265 epochs. Even though the denoising algorithm attenuates noise successfully, the reflection shape, which is important information of the SO data, can be altered. Therefore, it is necessary to measure the structural distortion to verify the effectiveness of the proposed method. The SSIM is a quality metric that calculates the structural similarity between two datasets and can be calculated as follows (Hore and Ziou, 2010):

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_u^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \qquad (5)$$

270 where μ is the average, σ^2 is the variance, σ_{xy} is the covariance of x and y, and c is a stabilizing parameter. The value of SSIM ranges from 0 to 1, and if the structure is distorted during the denoising process, the SSIM will be low. On the other hand, the SSIM will be close to 1 if the denoised data are similar to the ground truth. Fig. 10 (b) shows the average SSIM of the test results. Similar to the PSNR result, the SSIM is also low at the early stage of training but increases as training progresses and converges at approximately 0.88 after 20 epochs.

275 For seismic data, it is important to determine how well the actual amplitude and shape of the true reflection are recovered through the denoising process. In particular, the amplitude information is a key parameter for acquiring the data slope spectrum, which calculates slope spectra directly from the seismic data (Holbrook et al., 2013; Fortin et al., 2017). Therefore, we extracted seismic traces from the denoised section and ground truth and compared the extracted traces, as shown in Fig. 11, to ensure that the trained model recovers the actual amplitude of the signal. We extracted the 20th (Fig. 11 (a)) and 30th (Fig. 11 (b))

280 vertical traces from the last (6th) patch of the test data, as shown in Fig. 9. The amplitude and shape of the trace from the noisy data are different from those of the ground truth because the data are severely contaminated with noise. Even though there is much noise, the denoised traces have similar amplitudes and shapes to those of the ground truth. These results indicate that the DnCNN can recover important information of the true reflections and can be useful for random noise attenuation of sparker SO data.

285

Ocean Science
Discussions

Open Access

EGU

3.3 Experiment using training dataset 2

Training dataset 2 was generated by using the modified Marmousi-2 and Sigsbee2a synthetic seismic sections and noise obtained from the East Sea seismic section. After training the DnCNN model (D2 model) with training dataset 2, we evaluated the trained model against test data. The test data were generated with the same procedure of generating the training data, and

290   we selected the Static 94 synthetic seismic section, which is a different model than that used for the training data. The size of the test data patch was the same as that of the training data patch (50×50), and the number of test data points was 3,072. Fig. 12 shows 6 randomly selected test data, ground truth and denoised results after applying the D2 model at the 5th, 10th, 20th and 40th epoch. Even though the test data contain noise at different levels, the trained model at the 40th epoch attenuated most of the noise successfully and generated almost identical seismic sections to the ground truth. The second test data patch

295   contained relatively little noise, and most of the noise was removed after approximately 10 training epochs. Test data patches 1 and 3 contained simple reflections with much noise, and the noise was sufficiently removed after approximately 20 training epochs. The noise in test data patches 4 and 6 was more severe than the noise in the other test datasets. After 40 training epochs, most of the noise was attenuated but not perfectly removed. A weak signal remained in the bottom part of ground truth 5, and noise dominated test data patch 5. If we evaluate the denoised data, noise had been successfully removed, and only a weak

300   signal remained in the bottom part of the section after 40 training epochs. This indicates that the trained DnCNN model can accurately discriminate between signal and noise.

Unlike training dataset 1, training dataset 2 was generated with synthetic data. Therefore, it has the advantage of using noise-free seismic sections as the ground truth. In addition, generating many different kinds of synthetic seismic sections does not require much time or effort; thus, it is easy to increase the amount of training data compared to using field data as training data.

305   However, the features of synthetic seismic sections can be different from those of the target data requiring noise attenuation because the synthetic seismic sections were generated by simply convolving the reflection coefficient with the source wavelet. Several studies have applied machine learning to field seismic data interpretation, such as automated fault detection with synthetic training data (Wu et al., 2018), but machine-learning-based noise attenuation of SO data using synthetic training data has not yet been studied.

310   Similar to the previous experiment, we also calculated the average PSNR and SSIM to quantitatively verify the test results and compared the amplitudes of the extracted traces. Fig. 13 (a) shows the average PSNR, and 13(b) shows the average SSIM. Similar to the first experiment, the average PSNR and SSIM converged after approximately 25 epochs. However, the average PSNR and SSIM in the second experiment are higher than those in the first experiment. These results could be caused by the use of a noise-free synthetic seismic section as the ground truth of training dataset 2 and might indicate that training dataset 2

315   is more appropriate for random noise attenuation of SO data. Fig. 14 shows the extracted traces before and after applying the D2 model. We extracted the 20th (Fig. 14 (a)) and 30th (Fig. 14 (b)) vertical traces from the 1st patch of the test data, as shown in Fig. 12. The denoised traces successfully recovered the true amplitude and shape, although the input data were severely contaminated by random noise.

Ocean Science
Discussions

EGU
Open Access

In the second experiment, the noise-attenuated traces are closer to the ground truth traces than those in the first experiment.

320　However, the comparison of the several extracted traces does not indicate which training data are more suitable for suppressing noise of sparker SO data. Therefore, we calculated the root-mean-square (RMS) error between the denoised test data and ground truth of the test data and evaluated which training data produced a lower RMS error. The RMS error was calculated as follows:

$$RMS\ error = \sqrt{\frac{1}{ntest} \sum_{i=1}^{ntest} \sum_{j=1}^{nnode} (g_{ij} - d_{ij})^2} \qquad (6)$$

325　where g is the ground truth of the test data, d is the denoised test data, $ntest$ is the number of test data and $nnode$ is the size of each data point (50×50). Even though test datasets 1 and 2 were generated using the same noisy data (the part containing noise of the East Sea SO section), the RMS errors of test datasets 1 and 2 before noise attenuation were different, 6.374 and 6.339, respectively, because noise was randomly extracted from the noise data. Therefore, we normalized the RMS error by that of the test data before noise attenuation. Fig. 15 illustrates the normalized RMS error of the first and second experiments

330　at every epoch, and the normalized RMS errors were properly decreased in both results. The normalized errors converged at 0.268 in the first experiment and at 0.151 in the second experiment. The normalized RMS error of the second experiment is lower than that of the first experiment, indicating that the performance of the D2 model is better than that of the D1 model.

4 Application to the East Sea SO data

335   The DnCNN models trained with training datasets 1 and 2 (the D1 and D2 models, respectively) were applied to the East Sea SO data. We applied the trained DnCNN models to the seismic sections from 0.03 to 0.28 s (approximately 22.5 to 210 m) where the reflections exist, and Fig. 16 shows the processed East Sea water column seismic section (the processed seismic section) from 0 to 0.28 s. The seismic section shallower than 0.03 s is dominated by noise from direct waves, which is muted at the data processing stage, and the section deeper than 0.28 s mainly contains random noise. Fig. 17 (a) and (b) show the

340   seismic sections of lines 1 and 2, respectively, of the East Sea SO data after applying the D1 model. Most of the random noise was successfully removed, and the reflections became clearer. The strong random noise that occurred in the shallow part of the processed seismic sections was substantially attenuated, and the noise located between 150 and 200 km in the line 1 section and that between 220 and 270 km in the line 2 section were also properly removed. Since noise was successfully attenuated, reflections that were difficult to distinguish due to a low S/N ratio were clearly imaged. In particular, the weak signals of the

345   line 1 section between 0 and 50 km and between approximately 0.1 and 0.18 s became clearer after noise attenuation. In addition, the reflections with steep slopes between 25 and 50 km and between 0.12 and 0.2 s were obscured by severe noise, but the D1 model successfully attenuated the noise and clearly recovered the reflections. Fig. 17 (c) and (d) show the estimated noise of the line 1 and 2 sections with the D1 model. As shown, the D1 model successfully discriminated the noise component from the reflections; thus, the estimated noise sections are almost identical to the noise component of the processed seismic

350   section. Fig. 18 shows the result obtained by applying the D2 model. Fig. 18 (a) and (b) show denoised seismic sections, and (c) and (d) show the estimated noise of lines 1 and 2, respectively. Similar to the result from model D1 (Fig. 17), model D2 also successfully attenuated the random noise in the sparker water column seismic section. In addition, the estimated noise from the D2 model is almost similar to that from the D1 model.

    If we compare the results generated with the D1 and D2 models, they are similar. However, there are several differences.

355   Reflections are not observed from 150 to 200 km and at approximately 0.2 s in the line 1 seismic section. The result with the D1 model still contains noise in that part, while the result with the D2 model contains lower noise levels compared to that with the D1 model. In addition, for the weak reflections between 70 and 150 km and between 0.1 and 0.2 s, the reflections in the result from the D2 model are clearer and more continuous than those in the result from the D1 model. For line 2, from 220 to 260 km, noise can still be observed when the D1 model is applied, but most of the noise has been sufficiently suppressed when

360   the D2 model is applied. These results are probably because the SEZ data, which are the ground truth used to train the D1 model, are field data and contain noise to a certain degree. In other words, the D1 model is likely to regard the noise in the line 1 seismic section with similar characteristics to those contained in the ground truth as a signal rather than noise. On the other hand, the D2 model does not suffer from this kind of problem because its ground truth is noise-free synthetic data.

    The data slope spectrum is the slope spectrum calculated directly from seismic data and is an important parameter for SO

365   data analysis. Holbrook et al. (2013) suggested calculating the data slope spectrum before calculating the reflector slope spectrum because the random noise that should be removed before analyzing the seismic data becomes evident in the data

slope spectrum. Therefore, we calculated the data slope spectrum by using a part of the line 1 seismic section and compared the data slope spectra. Before calculating the data slope spectrum, we converted the seismic section from the time axis to the depth axis and extracted the part from 150 to 175 km and at a depth from 75 to 150 m. Fig. 19 (a), (b) and (c) show the seismic

370   sections extracted from the section before and after noise attenuation using models D1 and D2, respectively. The seismic section before noise attenuation was severely contaminated with random noise, but most of the noise was removed in the sections after noise attenuation. Fig. 19 (d) shows the calculated data slope spectra. From the KM07 model (Klymak and Moum, 2007), noise has a k_x^2 (horizontal wavenumber) slope in the slope spectrum, and we plotted the k_x^2 slope with the green dashed line in Fig. 19 (d) for comparison. The data slope spectrum of the section before noise attenuation has a k_x^2 at wavenumbers

375   above 0.002 cpm, which indicates that noise dominates these wavenumbers. Because of the severe noise, it is impossible to analyze the seismic data before noise attenuation. On the other hand, the data slope spectra after noise attenuation seem to contain internal waves subrange from 0.0015 to 0.006 cpm and turbulence subrange from 0.009 to 0.015 cpm that approximately follow $k_x^{-1/2}$ and $k_x^{1/3}$ slopes (Klymak and Moum, 2007), respectively. This result indicates that noise was properly attenuated and the seismic data could be analyzed, even though noise with a slope of k_x^2 still occurred at wavenumbers

380   above 0.02 cpm.

From the noise attenuation results obtained by applying the trained models to the East Sea sparker SO data, we showed that the DnCNN architecture used in this study can successfully suppress random noise. The comparison of the D1 and D2 model results showed that the training data generated using noise-free synthetic data are more suitable for random noise attenuation of sparker SO data than those generated using field data with a relatively high S/N ratio.

385

5 Conclusions

Random noise is one of the major obstacles in analyzing SO data. Conventionally, the noise in SO data has been attenuated through simple data processing methods because most of the SO data are obtained with air guns, which generates data with a relatively high S/N ratio. However, the simple noise attenuation method is not sufficient for data with a low S/N ratio, such as

390   sparker SO data. Therefore, we applied machine learning to attenuate the random noise in East Sea sparker SO data, which contains much random noise. The DnCNN architecture was used to construct the neural network, and training data were generated by combining the ground truth and noise extracted from the target seismic data at random amplitude ratios. Two different training datasets were generated, and they used either field or synthetic data as the ground truth. The trained DnCNN models were applied to the test datasets that were generated with the same procedure of generating the training datasets. The

395   test results were verified based on the PSNR, SSIM, trace extraction and normalized RMS error revealing that both trained DnCNN models were able to successfully attenuate random noise. We applied the trained DnCNN models to the East Sea sparker SO data, which is the target of this study, and the models successfully attenuated random noise. The comparison of the denoised seismic sections after applying the two different trained models showed that the training dataset generated from the noise-free synthetic data was more suitable for sparker SO data noise attenuation than that generated from the high-S/N ratio

400   field data. Even though the observed random noise is almost completely attenuated in the seismic section, the data slope spectrum still indicates that the section contains noise with a slope of k_x^2 at wavenumbers above 0.02 cpm. Therefore, future studies should include a detailed analysis of the slope spectra (both data and reflection slope spectra) of the East Sea SO data and establish an improved noise attenuation algorithm suitable for higher wavenumbers.

    The network architecture used in this study is straightforward and efficient. In addition, the proposed method of generating

405   the training dataset is very simple and easy because it only requires synthetic data, which are readily generated, and noise data, which can be extracted from the target seismic data. Moreover, only approximately one hour is required to train the DnCNN model with a single GPU. Therefore, the noise attenuation method suggested in this study has the advantage that it can be widely and easily applied in noise attenuation of the various kinds of SO data.

Ocean Science
Discussions

Open Access

EGU

410 **Data availability**

The synthetic training data can be downloaded from https://github.com/hgjun1026/so_dncnn.git. The field data and program

can be made available upon request to authors.

Author contribution

Hyunggu Jun and Hyeong-Tae Jou constructed machine learning program and performed experiments. Chung-Ho Kim, Sang

415 Hoon Lee and Han-Joon Kim acquired seismic data and performed data processing.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements.

420

Ocean Science
Discussions

Open Access

EGU

References

Araya-Polo, M., Farris, S., and Florez, M.: Deep learning-driven velocity model building workflow, The Leading Edge, 38(11), 872a1-872a9, 2019.

425

Blacic, T. M., Jun, H., Rosado, H., and Shin, C.: Smooth 2-D ocean sound speed from Laplace and Laplace-Fourier domain inversion of seismic oceanography data, Geophysical Research Letters, 43(3), 1211-1218, 2016.

Burger, H. C., Schuler, C. J., and Harmeling, S.: Image denoising: Can plain Neural Networks compete with BM3D?,
430   Computer Vision and Pattern Recognition 2012 IEEE Conference, 2392-2399, 2012.

Chen, Y., and Pock, T.: Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration, IEEE transactions on pattern analysis and machine intelligence, 39(6), 1256-1272, 2016.

435   Dagnino, D., Sallares, V., Biescas, B., and Ranero, C. R.: Fine-scale thermohaline ocean structure retrieved with 2-D prestack full-waveform inversion of multichannel seismic data: Application to the Gulf of Cadiz (SW Iberia), Journal of Geophysical Research: Oceans, 121(8), 5452-5469, 2016.

Dickinson, A., White, N. J., and Caulfield, C. P.: Spatial Variation of Diapycnal Diffusivity Estimated From Seismic Imaging
440   of Internal Wave Field, Gulf of Mexico, Journal of Geophysical Research: Oceans, 122(12), 9827-9854, 2017.

Fortin, W. F. J., Holbrook, W. S., and Schmitt, R. W.: Mapping turbulent diffusivity associated with oceanic internal lee waves offshore Costa Rica, Ocean Science, 12, 601–612, 2016.

445   Fortin, W. F., Holbrook, W. S., and Schmitt, R. W.: Seismic estimates of turbulent diffusivity and evidence of nonlinear internal wave forcing by geometric resonance in the South China Sea, Journal of Geophysical Research: Oceans, 122(10), 8063-8078, 2017.

Gondara, L.: Medical image denoising using convolutional denoising autoencoders. In: 2016 IEEE 16th International
450   Conference on Data Mining Workshops (ICDMW), 241-246, 2016.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778, 2016.

Ocean Science
Discussions

Open Access

EGU

455   Holbrook, W. S., Páramo, P., Pearse, S., and Schmitt, R. W.: Thermohaline fine structure in an oceanographic front from seismic reflection profiling, Science, 301(5634), 821-824, 2003.

Holbrook, W. S., Fer, I., Schmitt, R. W., Lizarralde, D., Klymak, J. M., Helfrich, L. C., and Kubichek, R.: Estimating oceanic turbulence dissipation from seismic images, Journal of Atmospheric and Oceanic Technology, 30(8), 1767-1788, 2013.
460

Hore, A., and Ziou, D.: Image quality metrics: PSNR vs. SSIM, In: 2010 20th International Conference on Pattern Recognition, 2366-2369, 2010.

Ioffe, S., and Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,
465   In: International Conference on Machine Learning, 448-456, 2015.

Jain, V., and Seung, S.: Natural image denoising with convolutional networks, In: Advances in Neural Information Processing Systems, 769-776, 2009.

470   Jun, H., Kim, Y., Shin, J., Shin, C., and Min, D. J.: Laplace-Fourier-domain elastic full-waveform inversion using time-domain modelling, Geophysics, 79(5), R195-R208, 2014.

Keras Documentation: https://keras.io/models/sequential/, last access: 20 March 2020.

475   Klymak, J. M., and Moum, J. N.: Oceanic isopycnal slope spectra. Part I: Internal waves, Journal of physical oceanography, 37(5), 1215-1231, 2007.

Krizhevsky, A., Sutskever, I., and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, In: Advances in neural information processing systems, 1097-1105, 2012.
480

Lefkimmiatis, S.: Non-local color image denoising with convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3587-3596, 2017.

Li, H., Yang, W., and Yong, X.: Deep learning for ground-roll noise attenuation, In: SEG Technical Program Expanded
485   Abstracts 2018, 1981-1985, 2018.

Liu, D., Wang, W., Chen, W., Wang, X., Zhou, Y., and Shi, Z.: Random noise suppression in seismic data: What can deep learning do?, In: SEG Technical Program Expanded Abstracts 2018, 2016-2020, 2018.

Ocean Science
Discussions

Open Access

EGU

490 Papenberg, C., Klaeschen, D., Krahmann, G., and Hobbs, R. W.: Ocean temperature and salinity inverted from combined hydrographic and seismic data, Geophysical Research Letters, 37(4), L04601, 2010.

Piété, H., Marié, L., Marsset, B., Thomas, Y., and Gutscher, M. A.: Seismic reflection imaging of shallow oceanographic structures, Journal of Geophysical Research: Oceans, 118(5), 2329-2344, 2013.

495

Ruddick, B. R.: Seismic Oceanography's Failure to Flourish: A Possible Solution, Journal of Geophysical Research: Oceans, 123(1), 4-7, 2018.

Sheen, K. L., White, N. J., and Hobbs, R. W.: Estimating mixing rates from seismic images of oceanic structure, Geophysical 500 Research Letters, 36(24), 2009.

Sheen, K. L., White, N. J., Caulfield, C. P., and Hobbs, R. W.: Seismic imaging of a large horizontal vortex at abyssal depths beneath the Sub-Antarctic Front, Nature Geoscience, 5(8), 542, 2012.

505 Si, X., and Yuan, Y.: Random noise attenuation based on residual learning of deep convolutional neural network, In: SEG Technical Program Expanded Abstracts 2018, 1986-1990, 2018.

Tang, Q., Hobbs, R., Zheng, C., Biescas, B., and Caiado, C.: Markov Chain Monte Carlo inversion of temperature and salinity structure of an internal solitary wave packet from marine seismic data, Journal of Geophysical Research: Oceans, 121(6), 510 3692-3709, 2016.

Tsuji, T., Noguchi, T., Niino, H., Matsuoka, T., Nakamura, Y., Tokuyama, H., Kuramoto, S. and Bangs, N.: Two-dimensional mapping of fine structures in the Kuroshio Current using seismic reflection data, Geophysical Research Letters, 32(14), 2005.

515 Wu, X., Liang, L., Shi, Y., and Fomel, S.: FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation, Geophysics, 84(3), IM35-IM45, 2019.

Wu, Y., DuBois, C., Zheng, A. X., and Ester, M.: Collaborative denoising auto-encoders for top-n recommender systems, In: Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, 153-162, 2016.

520

Xie, J., Xu, L., and Chen, E.: Image denoising and inpainting with deep neural networks, In: Advances in neural information processing systems, 341-349, 2012.

Ocean Science
Discussions

Yang, F., and Ma, J.: Deep-learning inversion: A next-generation seismic velocity model building method, Geophysics, 84(4),
525   R583-R599, 2019.

Zhang, K., Zuo, W., Chen, Y., Meng, D., and Zhang, L.: Beyond a gaussian denoiser: Residual learning of deep cnn for image
denoising, IEEE Transactions on Image Processing, 26(7), 3142-3155, 2017.

530   Zhao, X., Lu, P., Zhang, Y., Chen, J., and Li, X.: Swell-noise attenuation: A deep learning approach, The Leading Edge, 38(12),
934-942, 2019.

535

540

545

Figures
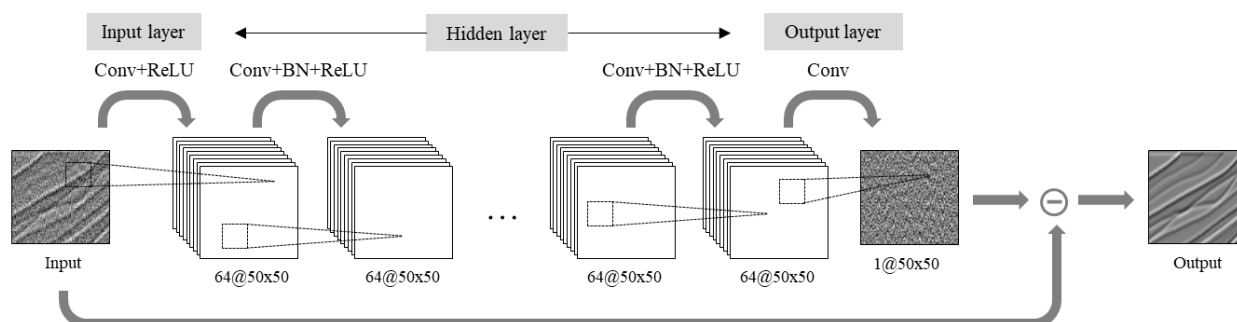
Figure 1: DnCNN architecture.

550
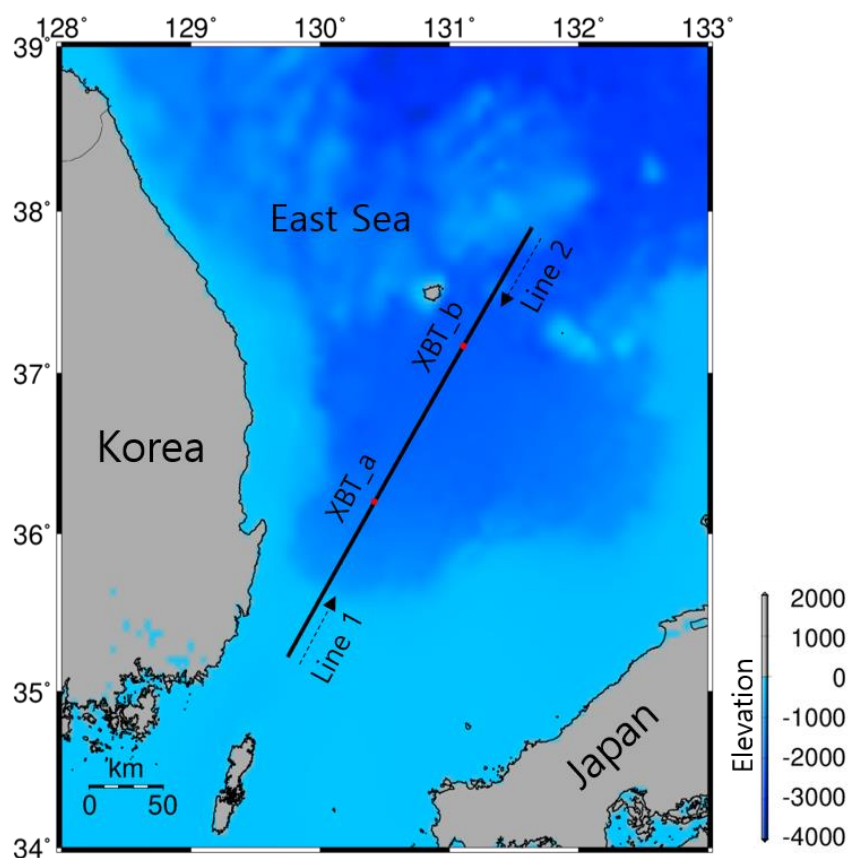
Figure 2: Location of seismic exploration. The black solid line is the survey line, and the black dashed lines indicate the exploration directions of lines 1 and 2.

Ocean Science
Discussions

Open Access

555

(a)

(b)
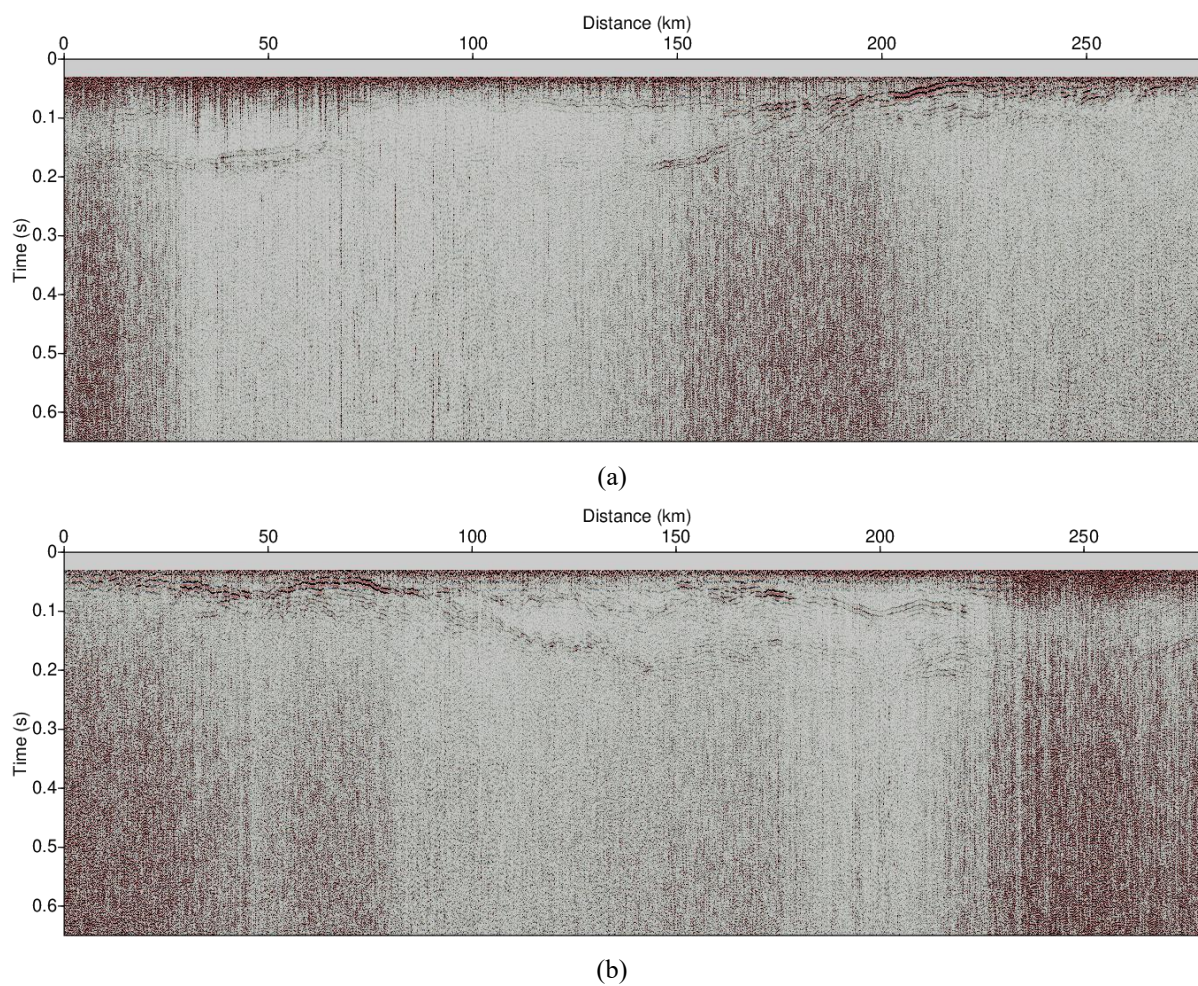
Figure 3: Processed seismic section of the East Sea: (a) line 1 and (b) line 2.

Ocean Science
Discussions
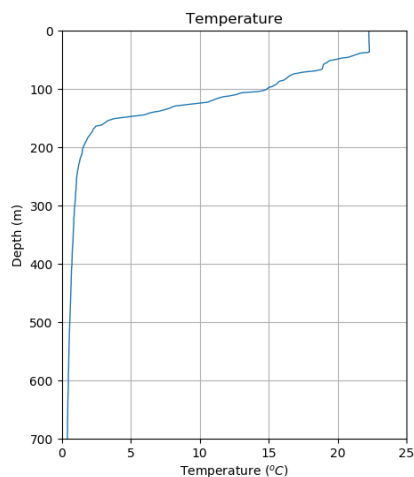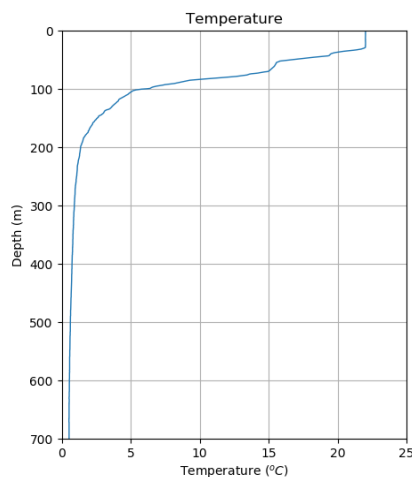
(a)
(b)

Figure 4: Temperature profiles obtained using XBT. (a) is XBT_a and (b) is XBT_b in Figure 3.
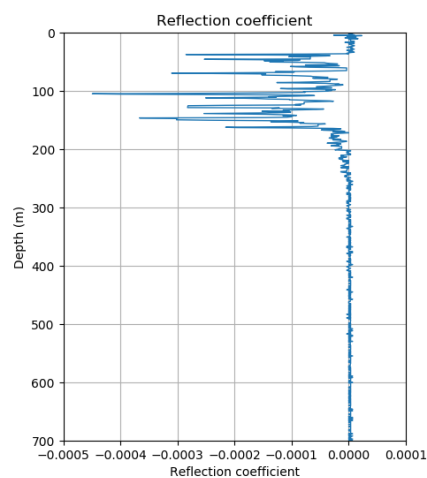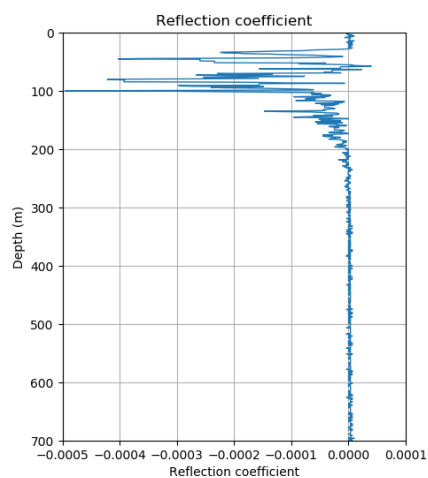
(a)
(b)

Figure 5: Reflection coefficients calculated using XBT data. (a) is XBT_a and (b) is XBT_b in Figure 3.

Ocean Science
Discussions

560

(a)

(b)
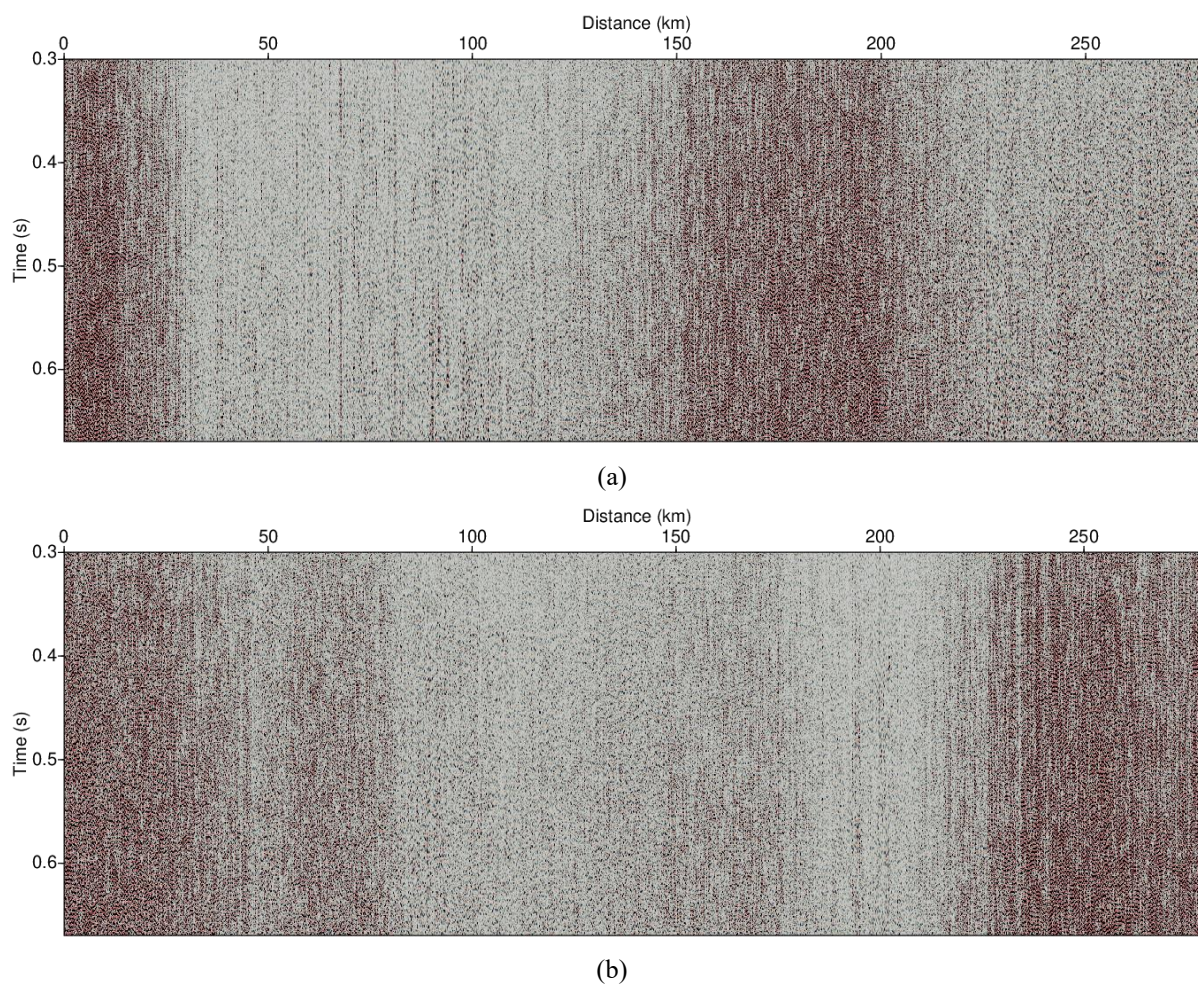
Figure 6: Noise section extracted from the East Sea SO data: (a) line 1 and (b) line 2.
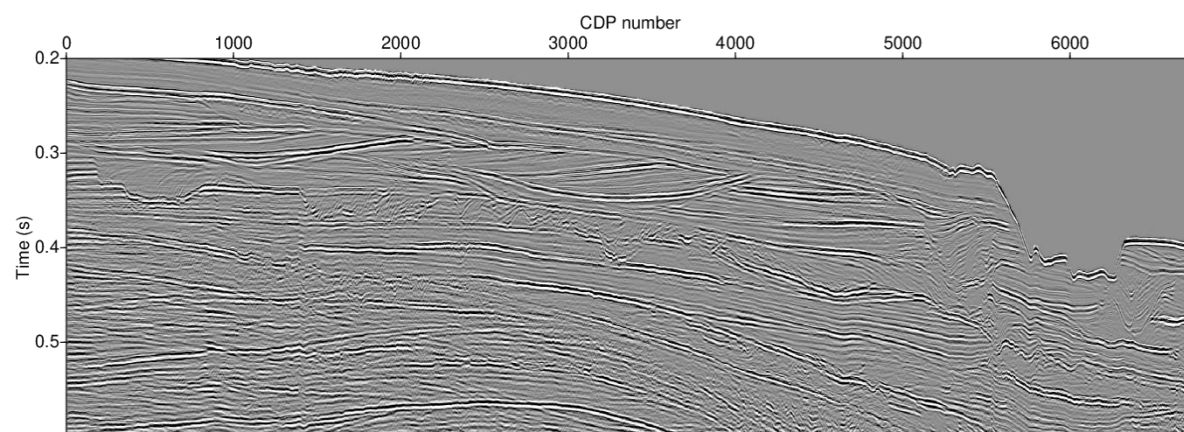
(a)
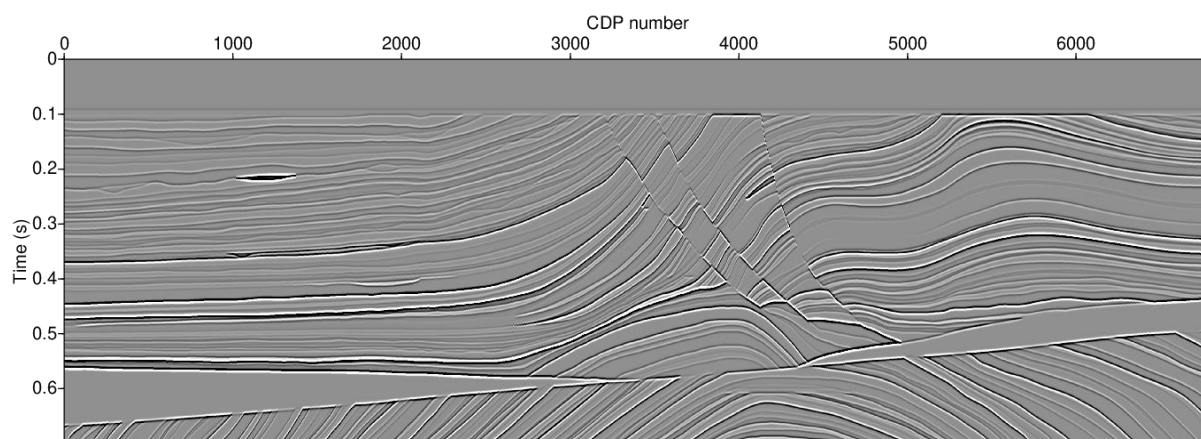
(b)
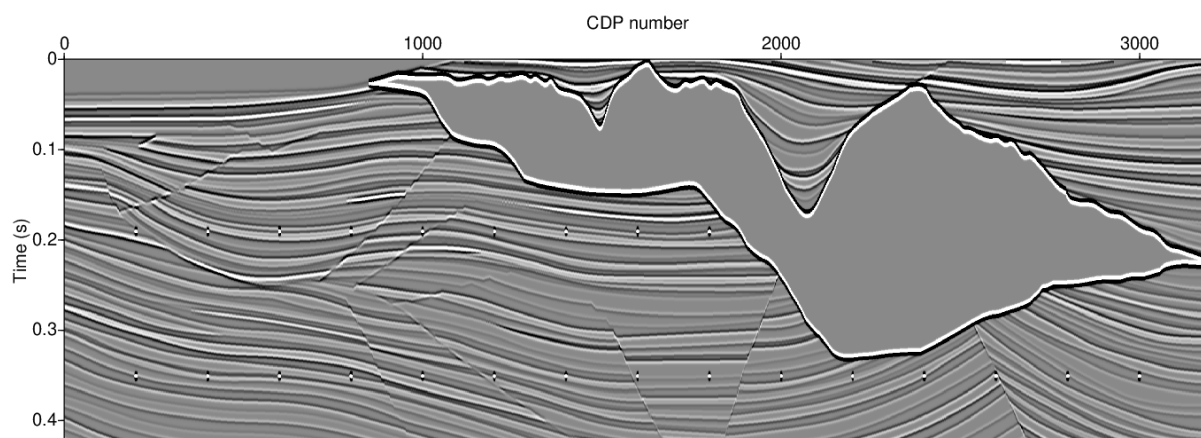
(c)

Figure 7: (a) Processed SEZ field seismic section, (b) Marmousi-2 synthetic seismic section and (c) Sigsbee 2A seismic section used to generate the training data.

Ocean Science
Discussions

Open Access

EGU
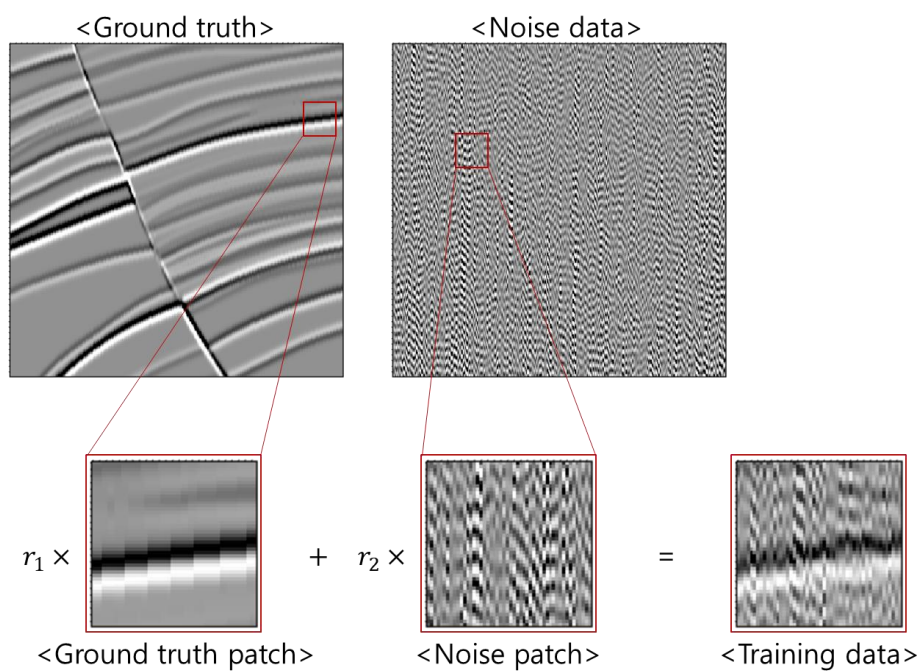
Figure 8: Example of constructing the training data.

565
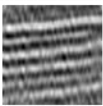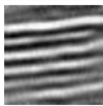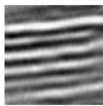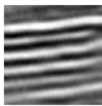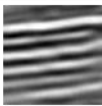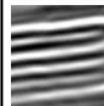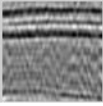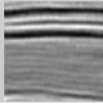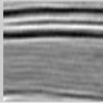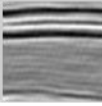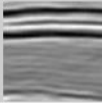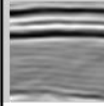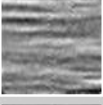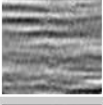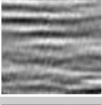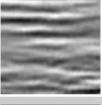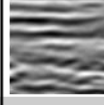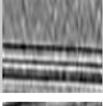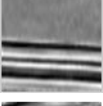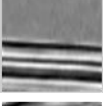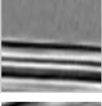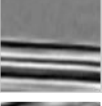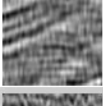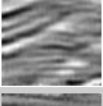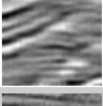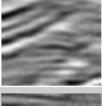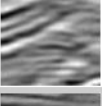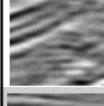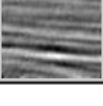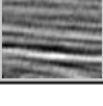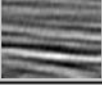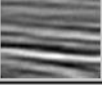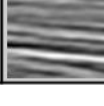
Figure 9: Test data, ground truth, and denoised results after applying the DnCNN models trained using training dataset 1.
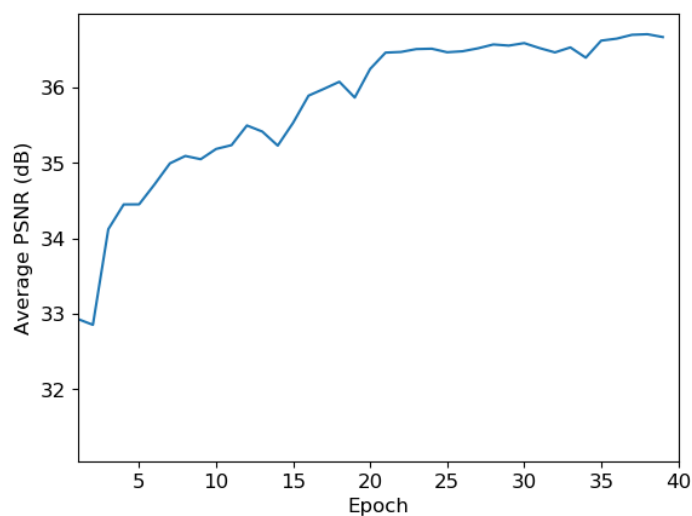
(a)
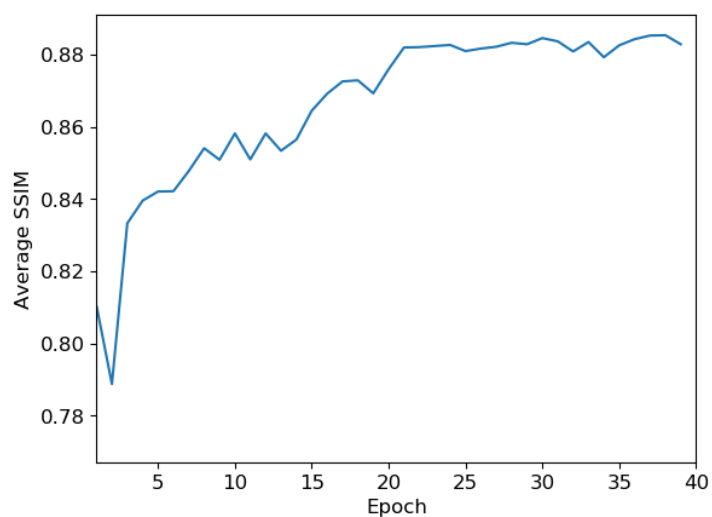
(b)

Figure 10: (a) PSNR and (b) SSIM of the test result of the first experiment.
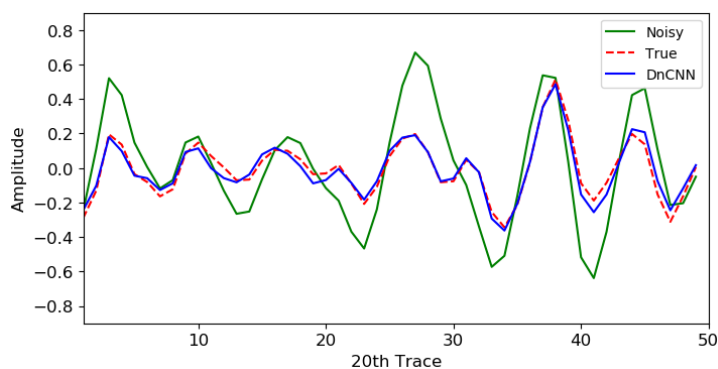
(a)
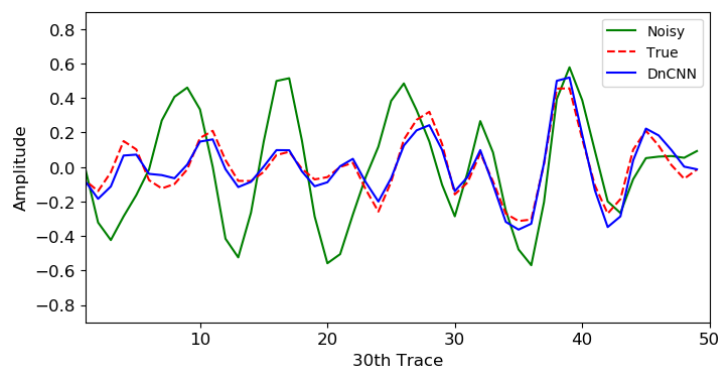
(b)

Figure 11: Comparison of the extracted traces before and after applying DnCNN using training dataset 1. The green solid line is the trace from the noisy data, the red dashed line is the trace from the ground truth and the blue solid line is the trace from the denoised data after applying the D1 model. (a) is the 20th and (b) is the 30th trace of the last test patch in Figure 10.

570

Figure 12: Test data, ground truth, and denoised results after applying the DnCNN models trained using training dataset 2.
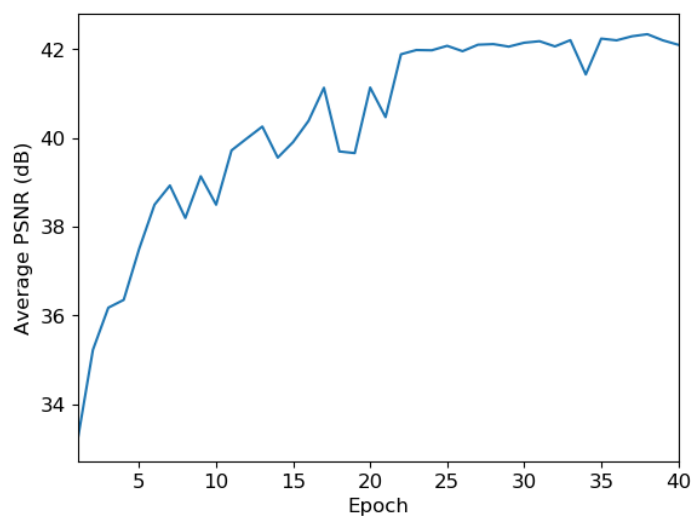
(a)
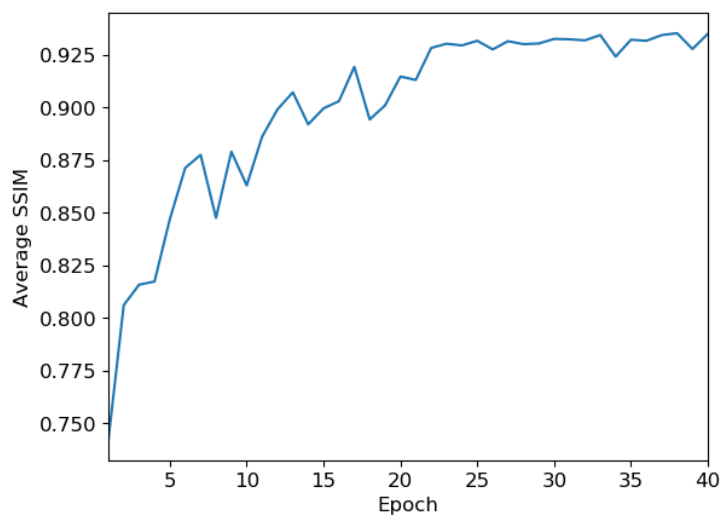
(b)

Figure 13: (a) PSNR and (b) SSIM of the test result of the second experiment.

575

(a)

(b)
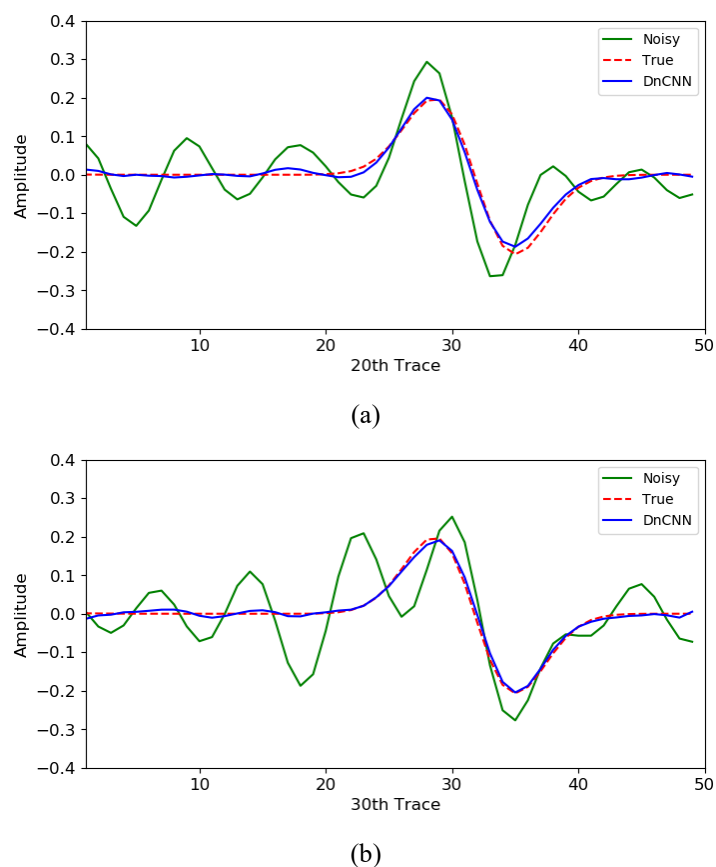
Figure 14: Comparison of the extracted traces before and after applying DnCNN using training dataset 2. The green solid line is the trace from the noisy data, the red dashed line is the trace from the ground truth and the blue solid line is the trace from the denoised data after applying the D1 model. (a) is the 20th and (b) is the 30th trace of the first test patch in Figure 13.
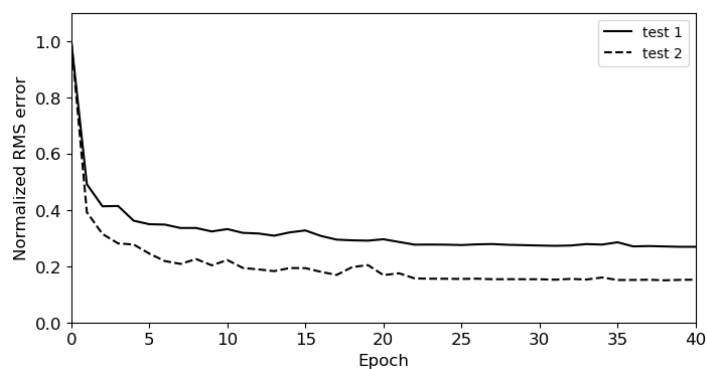
Figure 15: Normalized RMS error between the ground truth and denoised result of the first (solid) and second (dashed) experiments.

(a)

(b)
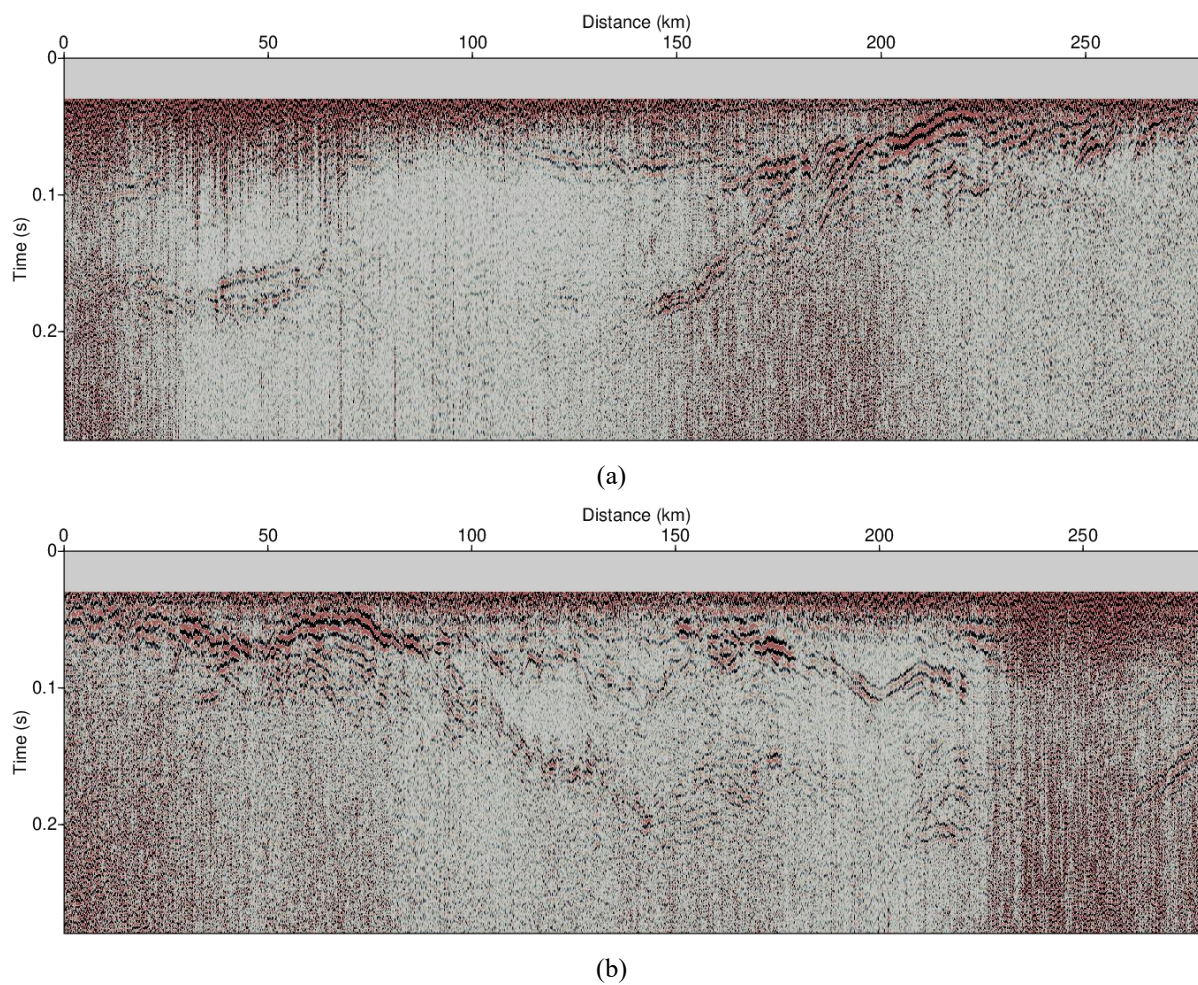
Figure 16: (a) Line 1 and (b) line 2 of the East Sea water column seismic section from 0 to 0.28 s.

580
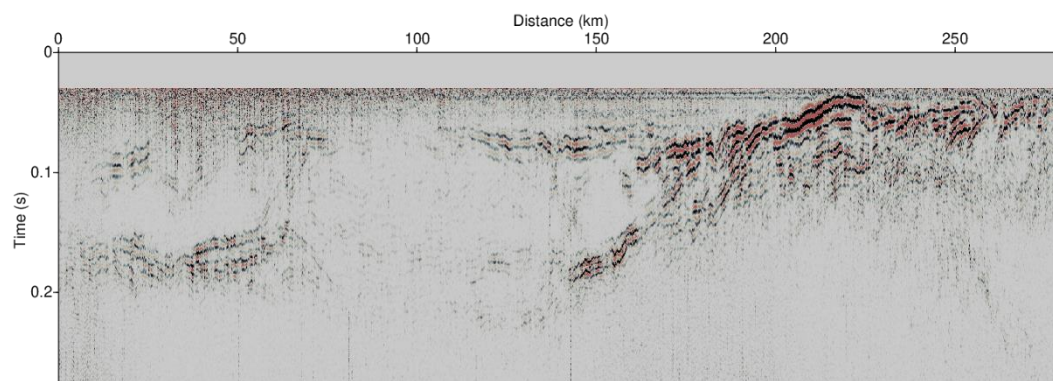
(a)
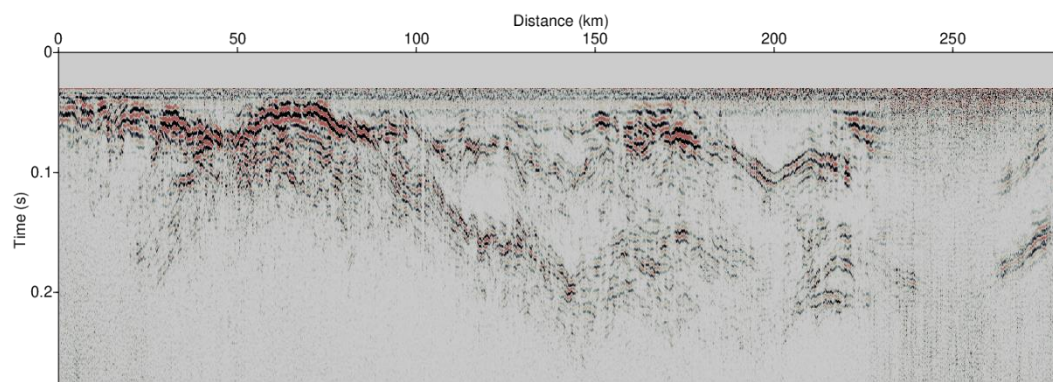
(b)
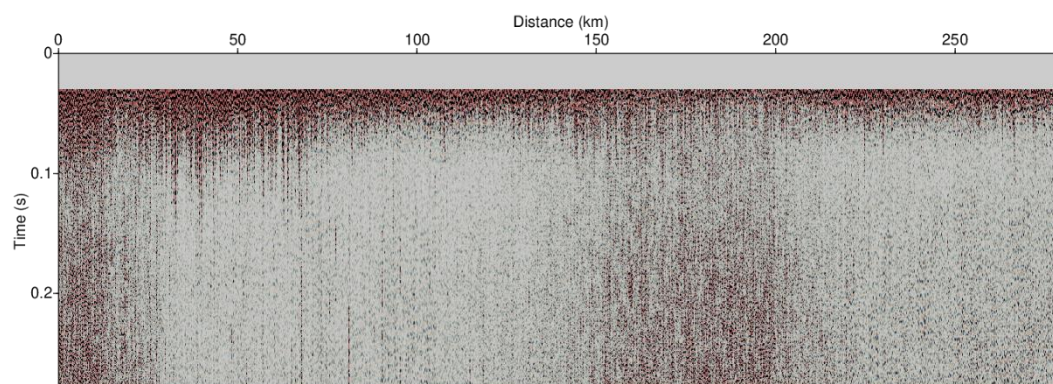
(c)
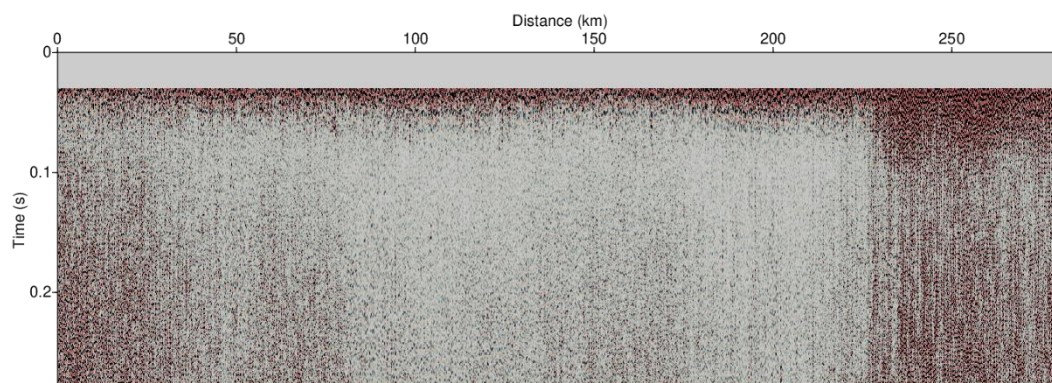
(d)

Figure 17: Noise attenuation results after applying the D1 model to (a) line 1 and (b) line 2 and estimated noise of (c) line 1 and (d) line 2.
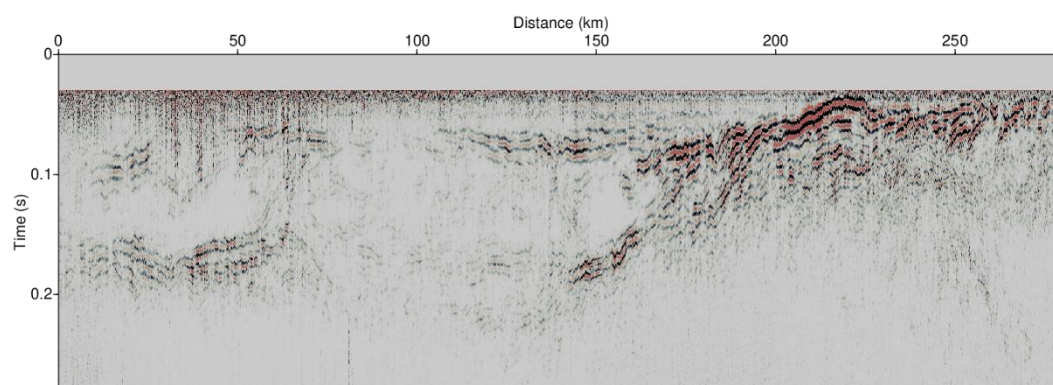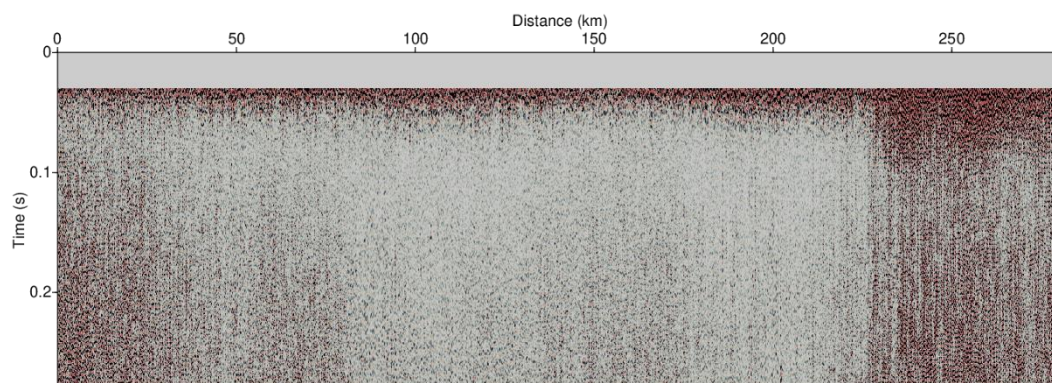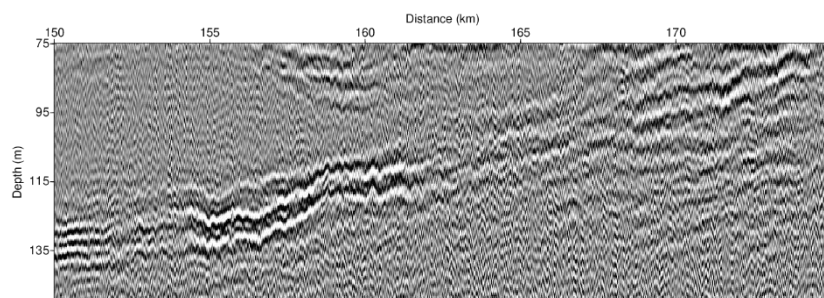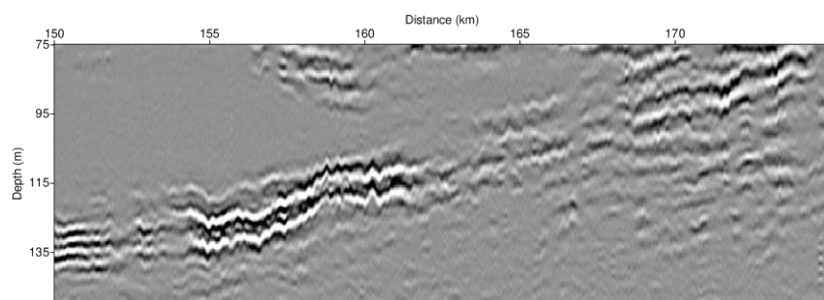
(a)
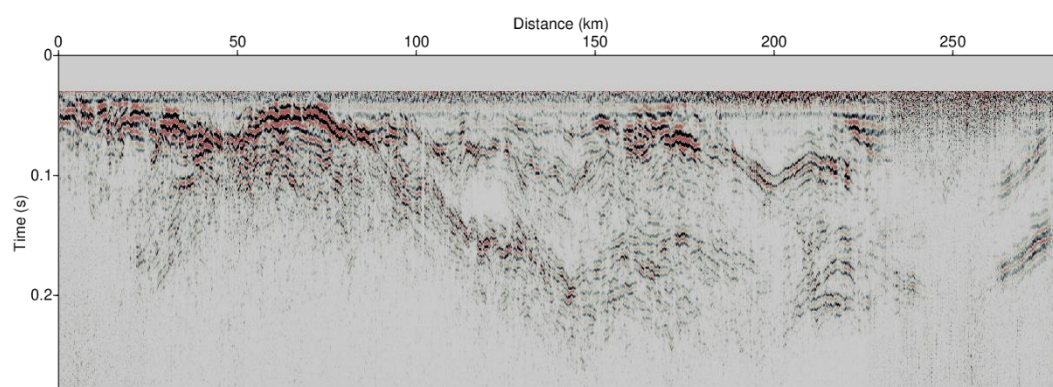
(b)
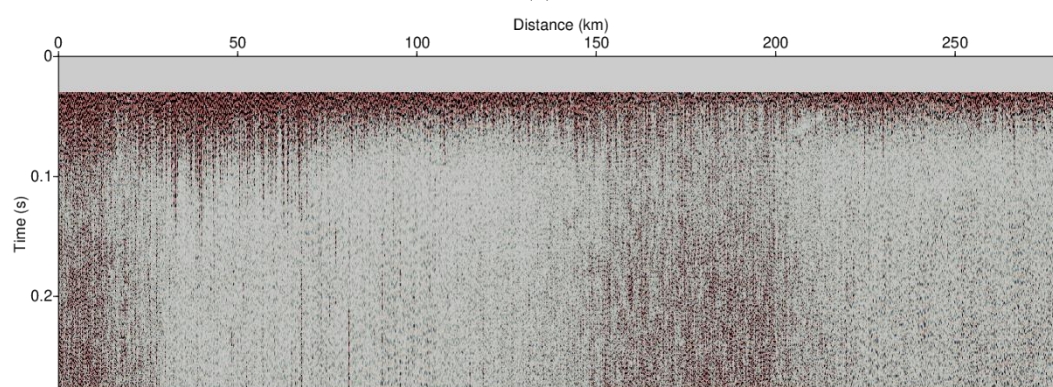
(c)

(d)

Figure 18: Noise attenuation results after applying the D2 model to (a) line 1 and (b) line 2 and estimated noise of (c) line 1 and (d) line 2.
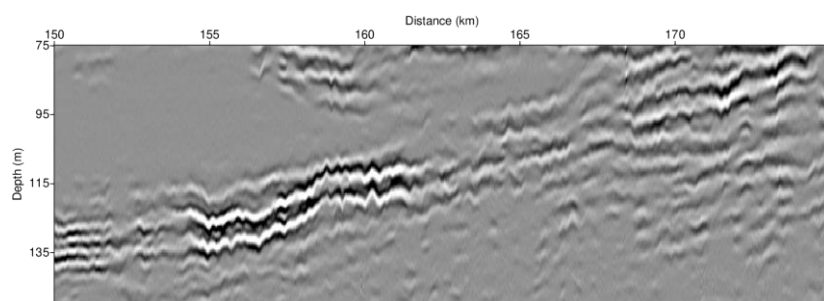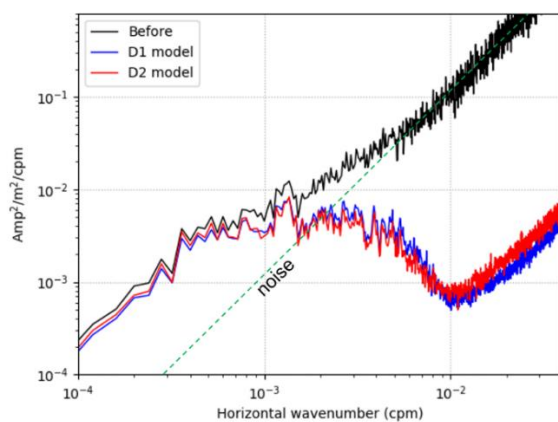
585

(a)

(b)

(c)

Ocean Science
Discussions

Open Access

EGU

(d)

Figure 19: Extracted seismic sections ((a) is the section before noise attenuation, (b) is the section after applying the D1 model and (c) is the section after applying the D2 model) and (d) shows calculated data slope spectra of (a), (b) and (c).

590