Reviewr 1

Thank you for your careful review and constructive comments. We have studied all of your comments carefully and revised our manuscript. Followings are the response to the Reviewer 1's comments. "**Q**" is a comment from the reviewer, "**A**" is a response to the comment and "**Changes**" the lines and details of the modification.

Q1. Li 6-7. The opening sentence of the abstract is a bit confusing. SO exploits water column reflections to interpret the oceanic features (fronts, eddies, water mass boundaries) as well as ocean fine structure (internal waves etc.). Futhermore, "compensating for the drawbacks of conventional PO equipment" is a very strong (and erroneous) statement. Perhaps "supplements the conventional PO observations". We should also be very careful when interpreting the seismic images to describe PO quantitatively.

A1. We modified the sentence.

Changes: Line 6-7. We modified the opening sentence to "Seismic oceanography (SO) acquires water column reflections using controlled source seismology and provides high lateral resolution that enables the tracking of the thermohaline structure of the oceans."

Q2. Li 8. The low / high frequency band introduction is not very helpful unless you relate it

to spatial resolution.

A2. We added the approximate frequency range with vertical resolution of each equipment.

Changes: Line 10-11. We added the resolution as "with a vertical resolution of approximately ten meters or more" and "with vertical resolution ranging from several centimeters to several meters."

Q3. Li 10. Reword "To solve the problem"? For example, "To extract reliable signal from the low S/N . . .."

A3. We modified the sentence.

Changes: Line 14. We modified the sentence as "To attenuate the random noise and extract reliable signal from."

Q4. Li 23: "measurements [from cruises] are performed. . .. . . observation [stations]."

A4. We modified the sentence.

Changes: Line 28. The modified sentence is "Conventional physical oceanography measurements from cruises are performed by dropping equipment at the observation stations"

Q5. Li 27: mention how the sea water characteristics can be estimated (through the acoustic impedance contrasts and expand a bit more to inform the reader)

A5. We added the explanation how the sea water is imaged

Changes: Line 32-34. The added sentence is "The differences in temperature and salinity between water column generate the difference in acoustic impedance, which reflect the seismic signals, and the reflected signals recorded at the receivers are processed to image the thermohaline fine structure of the ocean."

Q6, Q7. L29-32: Fine, but please do not oversell. Perhaps mention "qualitative images" and then move to "quantitative information after careful analysis where temperature/salinity contrasts produce well-defined horizons of seismic reflections" or similar. Also SO is not "widely used"

Li 32: reword "determine the behavior of turbulence and internal waves" to, for example, "quantify the internal wave spectral distribution and infer turbulence"

A6., A7. We modified the sentence.

Changes: Line 37-41. The modified sentence is "Therefore, SO is used to image the structure of water layers (Tsuji et al., 2005; Sheen et al., 2012; Piété et al., 2013; Moon et al., 2017) and provide quantitative information such as physical properties (i.e, temperature, salinity) (Papenberg et al., 2010; Blacic et al., 2016; Dagnino et al. 2016; Jun et al., 2019) or the spectral distribution of the internal wave and turbulence (Sheen et al., 2009; Holbrook et al., 2013; Fortin et al. 2016) after careful analysis where temperature or salinity contrasts produce clear seismic reflections"

Q8. Li 33: clarify what central frequency is (since the source covers a range of frequencies)

A8. We added the definition of the central frequency.

Changes: Line 44. We added "the geometric center of the frequency band (Wang, 2015)."

Q9. Li 43: vertical resolution of 1.5 m is not much superior to the vertical resolution of "several meters" stated in line 34. Perhaps specify the latter as 5-10 m?

A9. We modified specify the resolution.

Changes: Line 46. We modified "several meters" to "approximately ten meters or more."

Q10. Li 67-68: If not using MLP and AE (and any other acronym), no need to introduce them. It is difficult to read the text.

A10. We removed the explanation of MLP and AE.

Changes: Line 81-82. We removed the sentence "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)."

Q11. Li 64-75: If there's a possibility to thin out various methods introduced (and refer to a few key references and citations therein), it can be easier for the reader to follow.

A11. We thinned and removed several explanations.

Changes: Line 77-91. We rewrite this part as "The use of artificial intelligence (AI) has been studied in geophysics for decades (McComack, 1991; McCormack et al., 1993; Van der Baan and Jutten, 2000), but recent advances in computer resources and algorithms have spurred AI research, and several studies have 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. 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; 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. Recently, the DnCNN was applied to attenuate various types of noise in seismic data (Li et al., 2018; Si and Yuan, 2018; 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 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."

Q12. Li 77: East Sea appears very abruptly here, out of context.

A12. We removed the sentence of "East Sea".

Changes: Line 91-92. We removed the sentence "Therefore, this study applies the DnCNN to attenuate random noise in East Sea sparker SO data."

Q13. Li 135: delete "On the other hand,"?

A13. We removed "on the other hand"

Changes: Line 153. The modified sentence is "This study extracts noise from binary files, and a 3×3×1 convolution filter is adopted."

Q14. Sec 2.1 and 2.2: can any of these descriptions refer to Fig 1? (I only see a reference in the end, at li 141, and it is not very instructive.)

A14. We relocated the sentence

Changes: Line 145-148. The sentence "Fig. 1 shows the DnCNN architecture used in this study, where Conv and BN indicate convolution and batch normalization, respectively." is now located at the early part of the paragraph and matched explanation of each block to Fig. 1 like "This layer is shown as "Conv+ReLU" in Fig. 1."

Q15. Li 147: This is actually one line, but two repeats (in different travelling directions). Please mention the date of data collection, vessel speed during data collection. Transect duration etc.

A15. We added date of data collection with transect duration and vessel speed.

Changes: Line 167-169. The sentence "The survey was performed from October 7th to 11th in 2018 (approximately 38 hours for one line) and the vessel speed was 5.5 knots." is added.

Q16. Li 163-164: there're CTD /XCTD profiles, but the authors shown only 2xtemperature profiles from XBTs. It would be nice to increase the oceanographic context in the paper.

A16. We added 2 XCTD data.

Changes: Line 185. Temperature and reflection coefficients information from Two XCTDs are added in Fig. 4(a) and (b).

Q17. Li 167: please describe what a reflection coefficient is.

A17. We explained what a reflection coefficient is.

Changes: Line 187-188. The modified sentence is "Fig. 4 (b) shows the reflection coefficients, defining the ratio between the reflected and incident wave, calculated with the XBT and XCTD data."

Q18. Li 184-185: what do you mean by "thus, the subsurface seismic data have a better S/N ratio than the SO data."? Is subsurface seismic data not SO data? I suspect you mean beneath seabed by subsurface. Please clarify.

A18. We clarified the meaning.

Changes: Line 205, Line 212. We modified the “subsurface” to “below the sea floor and beneath seabed.” We also modified “sparker subsurface seismic data” to “SEZ seismic data”

Q19. Li 187: It is confusing: “We used the interval from 0.2 to 0.6 s of the original data where the noise level is relatively low”. Earlier you mentioned that part was just noise!

A19. There might be misunderstanding. The interval from 0.2 to 0.6 s of the SEZ data contains seismic data below the sea floor because the SEZ data is obtained shallow part of the East Sea where the water depth is approximately shallower than 200 m. However, the East Sea SO data which is the target data of this study is obtained from the deeper part of the East Sea and the water depth is approximately deeper than 1000 m. Therefore, the East Sea SO data contains random noise below 0.28 s (the water column) and SEZ data contains high S/N signal between 0.2 to 0.6 s (beneath the sea bed).

Q20. Li 190: Reword “the data are field data recorded with the same equipment.” as “the data are collected by the same equipment”

A20. We modified the sentence.

Changes: Line 213-215. The modified sentence is “This method has the advantage of using data with similar characteristics to those of the target data (the East Sea SO data) as the ground truth because the data are collected by the same equipment.”

Q21. Li 204: what is g/cc? Please use SI units.

A21. We changed unit to SI unit.

Changes: Line 228. We changed “g/cc” to “$1\ \mathrm{g/cm^3}$”

Q22. Li 249: bottom right (instead of right bottom)

A22. We modified “right bottom” to “bottom right”

Changes: Line 275.

Q23. Li 249-250: The sentence is confusing: “. . .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”. Dataset 1 has 6 test data. With the last reference to dataset 1 do you mean test data 1 or the entire dataset 1? Perhaps cut out the entire last part after the comma. Overall, I would appreciate a more distinct wording for test data. For example, subset 1 to 6, or patch (you use it in line 280)?

A23. We clarified the sentences.

Changes: Line 264-275. We modified “ground truth of test data 5” to “ground truth of the 5^{th} test data patch”. We also modified the “test data” to “test data patch”.

Q24. Li 279-280: 20th and 30th traces from the last patch: which epoch is this? Are the traces from the 50x50 patch? Can you please mention for the reader: “. . .traces out of the 50x50 size patch 6 of the test data”.

A24. We modified the sentence.

Changes: Line 308-310. The modified sentence is "We extracted the 20th (Fig. 10 (a)) and 30th (Fig. 10 (b)) vertical traces from the last (6th) patch of the test data, which had a size of 50x50. For the denoised trace, we extracted trace from the denoised patch of the 40th epoch."

Q25. Li 310-311: can be cut out; simply cross reference Fig 13 after 25 epochs. Overall there are repetitions throughout the authors could try to simplify.

A25. We simplified the paragraph by removing some repetitions.

Changes: Line 342-343. The sentence "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." was removed.

Q26. Li 325: perhaps specify, "is the number of test data patch (3072)"

A26. We modified the sentence.

Changes: Line 359. The sentence is modified as " is the number of test data patches (3,072)"

Q27. Li 327-328: too many significant digits at RMS errors? (perhaps enough with 6.37 and 6.34). For which epoch are these values? (Also the normalized values in line 331 could be 0.27 and 0.15)

A27-1. We reduced the significant digits.

A27-2. They are the RMS error of the test data before applying DnCNN. We modified the sentence to clarify the meaning.

Changes: Line 361-365. To clarify, we modified "RMS errors of test dataset 1 and 2 before noise attenuation…" to "initial RMS errors of test dataset 1 and 2 before noise attenuation…" We also changed 6.374, 6.339, 0.268, 0.151 to 6.37, 6.34, 0.27, 0.15.

Q28. Li 332: delete "than that of the D1 model"

A28. We removed "than that of the D1 model".

Changes: Line 366-367.

Q29. Eq 6, is a division by nmode missing?

A29. We wanted to calculate the average RMS error of each test data patch (not each node in a patch). Therefore, we divided the errors by ntest only.

Q30. Li 364: "The data slope spectrum is the slope spectrum. . ." this is all very confusing. The data slope spectrum is first referred to in line 276-277 (again without explanation). Please introduce what the data slope is. For example, "the slope spectrum is the horizontal wavenumber, k_x, spectrum of the horizontal gradient of the vertical displacement of a digitized horizon. The data slope spectrum is . . .?" (or a similar explanation. Note my interpretation of the slope spectrum can be in error.)

A30. To explain the data slope spectrum and avoid confusion we, we removed "data slope spectrum" in line 276-277 (in the original manuscript) which is unnecessary. Instead, we added the explanation of data slope spectrum at the synthetic data slope spectrum experiment part.

Changes: Line 305-306 and Line 368-374. We removed "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)." and added "Water column reflection data can be used to obtain the physical oceanographic information by calculating the slope spectrum. The data slope spectrum is a horizontal slope spectrum obtained directly from seismic data by calculating the horizontal wavenumber () spectrum of the seismic reflection amplitude, and it is useful to identify noise contamination of seismic data and the cutoffs from an internal wave to turbulence subrange (Holbrook et al., 2013; Fontin et al., 2017). 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."

Q31. Li 367: replace "we calculated the data slope spectrum . . .. and compared the data slope spectra" with "we calculated and compared the data slope spectra using the outcome of the D1 and D2 models. . ..

A31. We modified the sentence.

Changes: Line 444-445. We modified the sentence to "To validate the noise attenuation results, we also calculated and compared the data slope spectra by using the outcome of the D1 and D2 models."

Q32. Li 376: "slope" is missing before "at wavenumbers"

A32. We added "slope".

Changes: Line 454.

Q33. Li 377-378: I cannot quite follow the subranges and the mentioned slopes in this panel. Perhaps mark on the figure?

A33. We marked guide lines of each subrange in figure.

Changes: Line 457-458 and Fig. 20.

Q34. Li 390: Here again mention why sparker SO data may be desirable

A34. We mentioned why sparker SO data is desirable again.

Changes: Line 474-475. "Despite the low S/N problem, the sparker source has advantage of generating relatively high frequency band signal, which can provide information with higher vertical resolution."

Q35. Fig 1. Please offer some more explanation in the caption. If not possible, defer reader to the main text

A35. We added explanation of figure in caption of Fig. 1.

Changes: Fig. 1.

Q36. Fig 2. Elevation is grayed out for >0m, so the colorbar can stop at 0. It would be useful to add a few isobaths. I would call Line 1 and Line 2, Repeat 1 and Repeat 2.

A36. We added isobaths in the Fig. 2.

Changes: Fig. 2.

Q37. Figs 4 and 5 can be combined into 1 figure. I suggest two panels, T profiles in one panel with different color. Reflection coeff in the second panel with different colors and one offset by 0.0001 unit. Does the coefficient have a unit?

A37. We merged and modified figures, and the reflection coefficient does not have a unit.

Changes: Fig4.

Q38. Fig 6 can be removed. It is already shown in Fig 3 and with the statement time > 0.28 s. You can mark the region by a rectangle in Fig 3.

A38. We modified the Figure and marked the noise part by using red box.

Changes: Fig. 3.

Q39. Fig 10 (and Fig 13)- this is the average PSNR and SSIM for the 6 subsets of dataset 1? Would it not be better to show all 6 lines, or the average with one standard deviation? Actually the number of test data is 3072 (6 is an arbitrary pick), why not show the mean and std over all 3072? And also show a histogram?

A39-1. It is the average PSNR and SSIM of dataset 1 (Fig. 8) and dataset 2 (Fig. 12)

A39-2. We calculated standard deviation over all 3,072 test data and plotted.

Changes: Fig. 8 and 12.

Q40. Fig 11 caption: in the end the cross reference should be to Fig 9.

A40. We modified the cross reference.

Changes: Fig. 10.

Q41. Fig 14 caption: in the end the cross reference should be to Fig 12. In the text the model is referred to as D2 (but here D1).

A41. We modified the cross reference.

Changes: Fig. 14.

Q42. Fig 16. This figure is not needed either. It is simply the upper 0.28s of Fig 3. However, I appreciate that it is zoomed in and compared to the cleaned sections. See suggestion below Fig 17 comment.

Fig 17. Please consider removing xaxis labels from panels c to d, and placing panel labels in the upper left corner of panels (it's grayed out anyway), so that you can have a more condensed 4-panel, 1 page figure with minimum white space vertically between panels.

I think a reorganized version of Figs 16- 18 will be much better for the reader. I suggest 2 figures, each with 5 panels (with identical x-axis limits and width, and minimum white space between them). New Fig 16: Line 1 results. New Fig 17 Line 2 results with corresponding 5 panels for each figure:

a) data (as in Fig 16a) b) clean after D1 c) clean after D2 d) noise after D1 e) noise after D2.

A42. We re-ordered the figures. We also merged figures by removing unnecessary labels and white spaces.

Changes: Fig. 18 and 19.

General comment:

I would strongly encourage the authors to improve the "Ocean Science" part by adding more insight using the clean seismic sections. Can you make some oceanographic inferences, interpretations (or better quantification) using the cleaned Lines 1 and 2?

- ➔ This paper tries to apply machine learning technology to remove random noise from sparker SO data to help interpret SO data, and confirm the possibility of quantitative analysis. Because we want to focus on the noise attenuation method itself, the machine learning methodology and application of the proposed method are the main part of the paper. Adding the oceanographic analysis of denoised data will help authors understand the characteristic of the East Sea, but the manuscript will be vast in content. Therefore, after confirming the possibility of oceanographic analysis using denoised sparker SO data in this study, the detailed oceanographic analysis of East Sea data will be performed in the future study.

Finally, it would serve the community much better if the authors made available some code for noise attenuation using machine learning. They offer the code through communication with authors, but the impact would be far larger if they make the code available as a supplement.

- ➔ We also agree your comment. We think the distribution of the code is necessary for the community. However, the program will undergo some modification because the review process is not finished. After finishing the review process, we will distribute the program through github.

Reviewr 2

Thank you for your careful review and constructive comments. We have studied all of your comments carefully and revised our manuscript. We edited the English of the entire manuscript including “Abstract” by following Reviewer 2’s recommendation.

This paper deals with the noise attenuation method of sparker SO data using machine learning. The data obtained from the sparker source have advantages such as cheap data acquisition costs and high vertical resolution from several centimeters to several meters, but it has not been widely used in SO study and has not been quantitatively analyzed to date. This is mainly because of the low S/N ratio of the sparker seismic data. Due to strong noise, the conventional data processing method is not sufficient to attenuate the noise in the sparker seismic data, thus it is difficult to perform quantitative analysis such as calculating slope spectrum. Therefore, we would like to propose a method to suppress random noise in the sparker seismic data. This paper tries to apply machine learning technology to remove random noise from sparker SO data to help interpret SO data, and confirm the possibility of quantitative analysis. Because of this reason, the machine learning methodology and application of the proposed method are the main part of the paper. After confirming the possibility of oceanographic analysis using denoised sparker SO data in this study, the detailed oceanographic analysis of East Sea data will be performed in the future study. Followings are the response to the Reviewer 2’s comments. “**Q**” is a comment from the reviewer, “**A**” is a response to the comment and “**Changes**” the lines and details of the modification.

Q1. line 34 - delete "relatively low" as you do not state relative to what. Please edit paper to remove, as much as is possible, unqualified comparative statements.

A1. We removed “relatively low” in line 34 and removed unqualified comparative statements in the manuscript.

Changes: Line 9, 45, 52, 61, 211, 277 and 473.

Q2. line 46 - This problem is more accentuated in SO because the impedance contrasts between the layers are small.

A2. We modified the sentence.

Changes: Line 57-58. The modified sentence is “This problem is more accentuated in SO because the impedance contrasts between the water layers are smaller than the impedance contrasts between the layers beneath the seabed.”

Q3. line 65 - this reference list ignores the long history of the use of Neural Networks see McCormack’s paper in Leading Edge 1991 which shows an early attempt to use these NN to identify noisy traces in seismic data, since then NN have been evaluated for many tasks in the processing of seismic reflection data. Suggest authors change sense to recognise the history but equally highlight the recent advances in AI. I now note that this history is partly addressed in the following paragraph.

A3. We modified the sentence and added the history of Neural Network in seismic data processing.

Changes: Line 77-80. The modified sentence is “The use of artificial intelligence (AI) has been studied in geophysics for decades (McComack, 1991; McCormack et al., 1993; Van der Baan and Jutten, 2000), but recent advances in computer resources and algorithms have spurred AI research, and several studies have 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).”

Q4. line 153 - scaling by the sq-rt of time is not "spherical divergence" correction but a "geometric correction" as for true spherical divergence loss the amplitude scales by a 1/z which for a constant sound-speed medium is proportional to 1/t.

A4. We scaled the data by the sq-rt of time because we tried to make balance between noise in the shallow part which is affected by the tails of complex source wavelet and in the deep part of the data. We modified the phrase.

Changes: Line 173-174. We modified "spherical divergence correction" to "amplitude correction."

Q5. line 158 - an SVD filter can be effective in removing direct wave and maybe worth trying, though extreme care is needed to get offsets correct and correctly estimate of surface mixed layer sound-speed

A5. To remove the direct wave, SVD filter or Tau-P domain filter would be appropriate. However, the source signature of the sparker data is more complex than that of the air gun data, thus the filter may not properly eliminate the direct wave. Moreover, the noise near the sea surface is severe and the section before 0.03 second is not our research target (interesting SO signal does not exist in that part because this part is mixed layer which does not have large differences in reflection coefficient), therefore we muted the section before 0.03 second.

Q6. Fig 3 - plot sections in the same orientation and spatially lined up so it is possible to appreciate the similarity/differences in the two images but note in caption or by arrow on section the acquisition direction.

A6. We plotted the sections in the same orientation and added an arrow indicating the ship direction.

Changes: Fig 3.

Q7. line 183 - the subsurface will contain a range of reflection coefficients some will be tens to hundreds times larger but others will be of the order of magnitude as SO.

A7. We modified the sentence.

Changes: Line 205-207. The modified sentence is "The reflection coefficients of the major reflectors below the sea floor are tens to hundreds of times larger than that of the water column; thus, the seismic data below the sea floor have a better S/N ratio than the SO data."

Q8. lines 220-224 - definition of epochs and iterations is not clear.

A8. We modified the sentence to clarify the definition of epoch and iteration.

Changes: Line 247. The modified sentence is "The epoch is a process using all training data, and iteration is a process using a mini-batch; thus, an epoch usually consists of several iterations."

Q9. General question about noise - it is not clear, or I have overlooked the statement in the paper, but was the noise section extracted from data before or after divergence correction? If so, have you not imposed a time scaling on the noise as environmental noise levels would be expected to remain constant with time? So should this denoising be applied to non-divergence corrected data?

A9-1. We extracted the noise from the processed seismic section which was applied the amplitude correction. Even though the background noise level is supposed to be not influenced by the time, the noise level at the early time in the East Sea SO data is larger than the deep part of the section (this might be the noise related to the complex source wavelet of sparker). Therefore, we empirically selected square root of time as scaling factor to make balance of the noise amplitude from shallow to deep part of the section.

A9-2. Since we extracted the noise from the amplitude corrected seismic section, we applied the trained model to the amplitude corrected seismic section to remove the random noise. If we extracted the noise from non-amplitude-corrected data, then we should apply the trained model to the non-amplitude-corrected data.

A9-3. Before calculating the data slope spectra, we scaled the seismic section again by multiplying square root of time to each time step (consequently multiplying time to each time step of the data) for the spherical divergence correction.

Changes: Line 446-447. The modified sentence is "Before calculating the data slope spectrum, we scaled the seismic sections again by multiplying the square root of time to each time step (consequently multiplying the time to each time step) for the spherical divergence correction."

Q10. line 290 - what is the "Static 94 synthetic seismic section?

A10. We modified the sentence and added the reference.

Changes: Line 320. The sentence is changed to "part of the 1994 Amoco static test dataset (SEG Wiki)" and reference is "SEG Wiki: https://wiki.seg.org/wiki/1994_BP_statics_benchmark_model, last access: 22 June 2020."

Q11. Figs 16 & 17 see request for Fig 3.

A11. We modified the Figures.

Changes: Fig. 18 and 19.

Q12. A useful analysis would be to generate a synthetic with the expected spectral slopes then add noise at different levels and try to recover the input, the question I would like to know is is the shift after filtering (shown in Fig 19) removing weak signal too. Also discussion on the expected horizontal resolution. You state the peak frequency is 250 Hz which, after migration, should give a maximum horizontal resolution of ~1.5 m. However, it will be less as this is a 2D profile over a 3D structure so there will be out-of-plane contamination.

A12-1. We performed experiment using synthetic data.

Changes: Line 368-386 and Fig16. and 17. We explained the reason of performing the synthetic data experiment and showed the result of the experiment.

A12-2. We also can find the shifting of the spectrum between the noise added synthetic section and noise attenuated seismic sections at the wavenumber smaller than 0.001 cpm. However, the difference is also observed between the spectrum of noise free section and noise added section. In addition, the shifting is not observed between the spectrum of noise free section and noise attenuated section. Therefore, this shifting seems to be caused by the characteristic of the noise extracted from the East Sea SO data.

A12-3. We also mentioned the shift issue in the manuscript.

Changes: Line 459-464. The explanation of shift issue is "There is a shift in the data slope spectrum after noise attenuation at wavenumbers smaller than 0.001 cpm. This shift is also observed in the synthetic data slope spectrum experiments. In Fig. 17, there is a difference between the spectrum of the noise-added section and that of the noise-attenuated sections at wavenumbers smaller than 0.001 cpm. However, the difference is also observed between the spectrum of the noise-free section and that of the noise-added section. Therefore, this shift seems to be caused by the characteristic of the noise extracted from the East Sea SO data."

A12-4. In the conclusion, we added the limitation of 2D exploration related to the resolution. And we mentioned that it is necessary to acquire data by using 3D seismic exploration to improve the resolution

Changes: Line 489-496. The added sentence is "Even though the random noise is almost completely attenuated in the seismic section, the proposed method still needs several improvements. The observed random noise is successfully attenuated in the seismic section, but 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 of the East Sea SO data and establish an improved noise attenuation algorithm suitable for higher wavenumbers. Moreover, the data were collected using 2D seismic exploration, which can degrade the seismic resolution during the data processing stage because of the limitations of 2D seismic exploration such as out-of-plane contamination. Therefore, to improve the resolution of SO data, it is necessary to acquire data by using 3D seismic exploration.”

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

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

Abstract. Seismic oceanography (SO) acquires water column reflections using controlled source seismology and provides high lateral resolution that enables the tracking of the thermohaline structure of the oceans. ~~by seismic exploration compensating for the drawbacks of conventional physical oceanographic equipment.~~ Most SO studies obtain data using air guns, which can produce acoustic energy below 100 Hz bandwidth.~~have relatively low-frequency bands~~ with a vertical resolution of approximately ten meters or more. For higher-frequency bands, with vertical resolution ranging from several centimeters to several meters, ~~at a low exploration cost, using~~ a smaller, low-cost seismic exploration system may be used, such as a sparker source with central frequencies of 250 Hz or higher.~~a shorter receiver length, would be an alternative.~~ However, the sparker source has a relatively low energy compared to air guns and consequently produces data with a lower signal-to-noise (S/N) ratio. ~~To solve the problem~~ To attenuate the random noise and extract reliable signal from ~~of~~ the low S/N ratio of sparker SO data without distorting the true shape and amplitude of water column reflections, 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.~~ Specifically, we used ~~A~~ 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 training dataset. We ~~have~~ generated two different training datasets using synthetic and field data. Models trained with the different training datasets ~~are~~ were applied to the test data, and the denoised results ~~are~~ were quantitatively compared. To demonstrate the technique, the ~~The~~ trained models ~~are~~ were applied to an SO sparker seismic dataset acquired in the East Sea ~~the target seismic data, i.e., the East Sea sparker water column seismic reflection data,~~ and the denoised seismic sections ~~are~~ were evaluated. The results show that machine learning can successfully attenuate the random noise ~~of~~ in sparker water column seismic reflection data.

1 Introduction

Conventional physical oceanography measurements from cruises are performed by dropping equipment at ~~the~~ observation ~~points~~stations. 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~~, ~~. The differences in temperature and salinity between water column generate the difference in acoustic impedance, which reflect the seismic signals, and the reflected signals recorded at the receivers are processed to image the thermohaline fine structure of the ocean.~~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~~ image the structure of water layers (Tsuji et al., 2005; Sheen et al., 2012; Piété et al., 2013; Moon et al., 2017~~), ~~) and provide quantitative information such as physical properties (i.e, temperature, salinity) (Papenberg et al., 2010; Blacic et al., 2016; Dagnino et al. 2016; Jun et al., 2019) or the spectral distribution of the internal wave and turbulence (Sheen et al., 2009; Holbrook et al., 2013; Fortin et al. 2016) after careful analysis where temperature or salinity contrasts produce clear seismic reflections.~~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, the geometric center of the frequency band (Wang, 2015), 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 approximately ten meters or more~~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 is more accentuated in SO because the impedance contrasts between the water layers are smaller than the impedance contrasts between the layers beneath the seabed.~~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 influence of the background noise becomes ~~relatively~~ larger than when using a high-energy source. 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 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 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.

The use of artificial intelligence (AI) has been studied in geophysics for decades (McComack, 1991; McCormack et al., 1993; Van der Baan and Jutten, 2000), but ~~Recently~~recent~~, rapid~~ advances in computer resources and algorithms have spurred ~~artificial intelligence (~~AI~~)~~ research, and several studies have 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; 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. Recently, ~~T~~the DnCNN was applied to attenuate various types of noise ~~from~~ in 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 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 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 results are compared and evaluated.

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 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~~ in the hidden layer. A ~~L~~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. 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 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 DnCNN (C) can be expressed as follows:

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

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 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 mini-batch data distribution. The distribution of input data varies during training and the neural network has a risk of updating the weights ~~to~~ in 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~~ equal to 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 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. As mentioned above, ~~T~~the amplitude and shape of the seismic reflections are important for spectrum analysis using SO data. To minimize possible 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). Fig. 1 shows the DnCNN architecture used in this study, where Conv and BN indicate convolution and batch normalization, respectively. The first layer type consists of "convolution + rectified linear units (ReLUs; Krizhevsky et al. (2012))" and is used only at the first layer of the network architecture. This layer is shown as "Conv+ReLU" in Fig. 1.~~.~~ In the convolution process, 2-dimensional convolution between ~~the~~ a certain size of kernel and data is performed. The outputs of the convolution process are passed 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.), ~~and~~ thus 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, t~~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. This layer is shown as "Conv+BN+ReLU" in Fig. 1. Sixty-four $3 \times 3 \times 64$ convolution filters are used because the number of feature maps of the hidden layer is 64, which is the same ~~at~~ for all hidden layers. After convolution, batch normalization and the ReLU activation function are applied. The third layer type is convolution and uses only the last layer to generate output noise data, and one $3 \times 3 \times 64$ convolution filter is used. This layer is shown as "Conv" in Fig. 1. 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

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 survey was performed from October 7th to 11th in 2018 (approximately 38 hours for one line) and the vessel speed was 5.5 knots. 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 from these sections did not target the layers below the sea floor~~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~~ amplitude corrections, bandpass filtering, common-midpoint (CMP) sorting and stacking. ~~Spherical divergence~~ Amplitude correction was performed by empirically multiplying the square root of time at each time step. The corner frequencies of the trapezoidal bandpass filter (Dickinson et al. 2017) ~~was~~ were 60-80-250-300 Hz, which ~~was~~ were higher than ~~that~~ those 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 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 (a) shows the temperature profiles from two XBTs and two XCTDs ~~casting 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 approximately from 30 to 200 m and deep water occurred below approximately 200 m depth. Fig. ~~5~~ 4 (b) shows the reflection coefficients, defining the ratio between the reflected and incident wave, calculated with the XBT and XCTD data ~~and~~ (assuming a constant density 1 g/cm^3). 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 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 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 noise parts of the East Sea SO data are shown as red boxes in Fig 3. ~~The sections with noise extracted from the East Sea SO data are shown in Fig. 6.~~ There are no notable reflections in the ~~sections with noise~~noise parts. 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 below the sea floor. The reflection coefficients of the major reflectors below the sea floor ~~of the subsurface~~ ~~is~~ are tens to hundreds of times larger than that of the water column; thus, the ~~subsurface~~ seismic data below the sea floor have a better S/N ratio than the SO data. In addition, after the proper data processing steps, the S/N ratio of the ~~subsurface~~ seismic data beneath the sea bed can be further enhanced. We used 14 lines of ~~subsurface~~ field sparker seismic data targeting below the sea floor (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 lower than in other parts of the 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~~ 5 (a) shows an example of the SEZ~~sparker subsurface seismic~~ data used to generate training dataset 1. This method has the advantage of using data with similar characteristics to those of the target data (the East Sea SO data) as the ground truth because the data are collected by ~~field data recorded with~~ the same equipment. Even if the S/N ratio of the sparker ~~subsurface~~ seismic data beneath the sea bed is relatively higher than that of the sparker ~~SO data~~seismic data of the water column 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.

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 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 ($1\ \mathrm{g/cm^3}$~~1 g/cc~~) were employed to calculate the reflection coefficient, 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~~ 5 (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 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~~ 6 shows an example training data compilation.

$$T = r_1 \times G + r_2 \times N, \tag{3}$$

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~~2~0.8 ($r_1 + r_2 = 1$). 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 number of ground truth data ~~size~~, data augmentation was applied by zooming in/out and randomly rotating or flipping the 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~~ patches at every epoch because the mini-batch size was 128 and the iteration of each epoch was 220. The epoch is a process ~~of~~using all training data, and iteration is a process ~~of~~using a mini-batch; thus, an epoch usually consists of several iterations. ~~means the number of processes using entire data point and iteration means the number of processes using a mini-batch in a epoch.~~

3 Training

3.1 Experimental setting

The experiment was conducted using 28,160 training data ~~points~~ patches 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.'s (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~~ that were not used to generate the training data. ~~The number of~~Eighty-six 300 × 300 size test data ~~was 86 and~~were available, 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~~ 7 shows 6 randomly selected test data subset patches, ground truths and denoised results after applying the D1 model at the 5^{th} , 10^{th} , 20^{th} and 40^{th} epochs. The depicted test data patches (1 to 6) include noise, but most of the noise has been successfully removed after training for 40 epochs. ~~Especially in test data~~In the 3^{rd} and 6^{th} patches of test data subset in particular, 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 the 4^{th} test data patch~~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~~ the 5^{th} test data patch contains noise in the ~~right~~ bottom right part, and training dataset 1 might also contain~~s~~ 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 below the sea floor~~, which had a relatively high S/N ratio~~, noise still 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~~ 5 (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) by using entire test data. The PSNR reflects the amount of noise contained in the data and can be calculated as follows (Hore and Ziou, 2010):

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

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~~ 8 (a) shows the average PSNR and standard deviation 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 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):

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

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~~ 8 (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~~ 21 epochs. We also plotted the PSNR and SSIM histogram of the test data before and after applying the D1 model (40th epoch) in Fig. 9. Both the PSNR and SSIM are clearly improved after applying D1 model.

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~~10, to ensure that the trained model recovers the actual amplitude of the signal. We extracted the 20th (Fig. ~~11~~ 10 (a)) and 30th (Fig. ~~11~~ 10 (b)) vertical traces from the last (6th) patch of the test data, which had a size of 50x50.~~as shown in Fig. 9.~~ For the denoised trace, we extracted trace from the denoised patch of the 40th epoch. 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.

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~~ used to generate~~ing~~ the training data, and we selected ~~the~~ part of the 1994 Amoco static test dataset (SEG Wiki, 2020)~~Static 94 synthetic seismic section~~, which is a different model from ~~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~~ 11 shows 6 randomly selected test data subset patches, ground truths and denoised results after applying the D2 model at the 5^{th}~~-~~, 10^{th} , 20^{th} and 40^{th} epochs. Even though the test data patches contain noise at different levels, the trained model at the 40^{th} epoch attenuated most of the noise successfully and generated almost identical seismic sections to the ground truth. The second test data patch contained relatively little noise compared to other test data patches, 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~~data patches. After 40 training epochs, most of the noise was attenuated but not perfectly removed. The noise was dominant in test data patch 5, and only a weak signal existed~~ A weak signal remained~~ in the bottom part of ground truth 5.~~, and noise dominated test data patch 5.~~ If we evaluate the denoised result of the 5^{th} test data patch~~data~~, noise had been successfully removed, and only a weak signal remained in the bottom part of the ~~section~~ patch 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. 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~~ by training the model using synthetic data, 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.

~~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 (Fig. 12 (a)) and SSIM (Fig. 12 (b)) converged after approximately 25 epochs. The histograms of PSNR and SSIM of the test data before and after applying the D2 model (40th epoch) are also plotted in Fig. 13. As shown, the PSNR and SSIM are improved after DnCNN is applied. ~~However, t~~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 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.

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

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

where g is the ground truth of the test data, d is the denoised test data, $ntest$ is the number of test data patches (3,072) 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 initial RMS errors of test datasets 1 and 2 before noise attenuation were different, 6.374 and 6.3~~39~~4, 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 at every epoch, and the normalized RMS errors were properly decreased in both results. The normalized errors converged at 0.~~268~~ 27 in the first experiment and at 0.15~~1~~ 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~~.

Water column reflection data can be used to obtain the physical oceanographic information by calculating the slope spectrum. The data slope spectrum is a horizontal slope spectrum obtained directly from seismic data by calculating the horizontal wavenumber (k_x) spectrum of the seismic reflection amplitude, and it is useful to identify noise contamination of seismic data and the cutoffs from an internal wave to turbulence subrange (Holbrook et al., 2013; Fontin et al., 2017). 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 and compared the data slope spectrum of noise-free, noise-added and noise-attenuated seismic data by using synthetic seismic section to verify that the proposed denoising method can recover the true data slope spectrum. The synthetic seismic section was generated by convolving the source wavelet with a randomly generated reflection coefficient section. Then, the noise extracted from the East sea SO data was added. Fig. 16 (a) shows the generated synthetic water column reflection section, and Fig. 16 (b) shows the noise added section. We applied the trained D1 model and D2 model to attenuate the noise, and the results are in Fig. 16 (c) (D1 model) and (d) (D2 model). Most of the noise was successfully attenuated, but the noise was not perfectly removed in the D1 model result at a distance from 20 to 25 km and depth from 140 to 180 m. Fig. 17 shows the calculated data slope spectra. The data slope spectrum of the noise-added section follows a k_x^2 slope, which is the slope of the random noise. After the noise attenuation, the data slope spectrum of the D2 model result (red line) follows the data slope spectrum of the noise-free section (greed line) almost identically. The data slope spectrum of the D1 model result (blue line) does not follow the noise slope, but the data slope spectrum is distorted compared to the noise-free data. The comparison of data slope spectra using synthetic data shows that the D2 model can recover the true data slope spectrum better than the D1 model.

4 Application to the East Sea SO data

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. 18 shows the results of applying the DnCNN to line 1. Fig 18 (a) is the line 1 seismic section from 0 to 0.28 s before the noise attenuation. 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. 18 (b) and (c) are the denoised seismic section after applying the D1 and D2 model, respectively. ~~Fig. 17 (a) and (b) show the seismic sections of lines 1 and 2, respectively, of the East Sea SO data after applying the D1 model.~~ In both results, ~~M~~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 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 t~~ Fig. 18 (d) and (e) are the estimated noise using the D1 and D2 model, respectively. ~~he estimated noise of the line 1 and 2 sections with the D1 model.~~ As shown, ~~the D1~~both models successfully discriminated the noise component from the reflections; thus, the estimated noise sections are almost identical to the noise component of the processed seismic section. Even though both models successfully attenuated the noise in the seismic section of line 1, there are several differences. Reflections are not observed from 150 to 200 km and at approximately 0.2 s in the line 1 seismic section. The result from the D1 model still contains noise in that part, while the result from the D2 model contains lower noise levels compared to that from 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.

Fig. 19 shows the results of applying the DnCNN to line 2. Fig .19 (a) shows the line 2 seismic section from 0 to 0.28 s before the noise attenuation. Fig. 19 (b) and (c) show the denoised seismic section after applying the D1 and D2 model, respectively. The seismic section of line 2 was contaminated by severe noise, but the D1 and D2 model properly removed the noise. In particular, the strong random noise located between 0 to 50 km was removed; thus, it became possible to recognize the reflections that were illegible. In addition, the reflections with steep slopes between 240 and 260 km and between 0.12 and 0.2 s were obscured by severe noise, but the D1 and D2 models successfully attenuated the noise and clearly recovered the reflections. However, similar to the line 1 result, the D2 model attenuated the noise better than the D1 model in some parts of the section. From 20 to 50 km and 250 to 280 km, noise can still be observed when the D1 model is applied, but most of the noise has been sufficiently suppressed when the D2 model is applied.

~~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. 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 the D2 model is applied.~~ Despite the successful noise attenuation of the D1 and D2 models, we found some differences. We presume ~~T~~that these ~~results~~ differences are caused by the characteristics of the SEZ data which are the ground truth used to train the D1 model. The SEZ data are field data and contain noise to a certain degree because it is almost impossible to perfectly remove the noise from the field data. ~~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 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, w~~To validate the noise attenuation results, we also calculated and compared the data slope ~~spectrum~~ spectra by using the outcome of the D1 and D2 models. ~~a part of the line 1 seismic section and compared the data slope spectra.~~ Before calculating the data slope spectrum, we scaled the seismic sections again by multiplying the square root of time to each time step (consequently multiplying the time to each time step) for the spherical divergence correction. Then, 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~~ 20 (a), (b) and (c) show the seismic 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~~ 20 (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~~ 20 (d) for comparison. The data slope spectrum of the section before noise attenuation has a k_x^2 slope at wavenumbers 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 the $k_x^{-1/2}$ (yellow dashed line) and $k_x^{1/3}$ (purple dashed line) 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 above 0.02 cpm. There is a shift in the data slope spectrum after noise attenuation at wavenumbers smaller than 0.001 cpm. This shift is also observed in the synthetic data slope spectrum experiments. In Fig. 17, there is a difference between the spectrum of the noise-added section and that of the noise-attenuated sections at wavenumbers smaller than 0.001 cpm. However, the difference is also observed between the spectrum of the noise-free section and that of the noise-added section. Therefore, this shift seems to be caused by the characteristic of the noise extracted from the East Sea SO data.

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.

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 sparker SO data. Despite the low S/N problem, the sparker source has advantage of generating a higher-frequency band signal than an air gun source, which can provide information with higher vertical resolution. 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 test results were verified based on the PSNR, SSIM, trace extraction and normalized RMS error. The data slope spectrum test using synthetic seismic section was also performed. The test results revealed that both trained DnCNN models were able to successfully attenuate random noise and the training data generated using noise-free synthetic data showed better results than the training data generated using high-S/N ratio field data. ~~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 also 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 field data.

Even though the ~~observed~~ random noise is almost completely attenuated in the seismic section, the proposed method still needs several improvements. The observed random noise is successfully attenuated in the seismic section, but 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 of the East Sea SO data and establish an improved noise attenuation algorithm suitable for higher wavenumbers. Moreover, the data were collected using 2D seismic exploration, which can degrade the seismic resolution during the data processing stage because of the limitations of 2D seismic exploration such as out-of-plane contamination. Therefore, to improve the resolution of SO data, it is necessary to acquire data by using 3D seismic exploration. ~~the data slope spectrum still indicates that the section contains noise with a slope of~~ $\cancel{k_x^2}$ ~~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 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.

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 the machine learning program and performed experiments. Chung-Ho Kim, Sang 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.

This research is supported by the Korea Institute of Ocean Science and Technology (Grant number PE99841, PE99851) and the Korea Institute of Marine Science and Technology promotion (Grant number 20160247).

References

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

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?, 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.

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 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.

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 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.

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.

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, 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.

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.

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.

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 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.

McCormack, M. D.: Neural computing in geophysics, The Leading Edge, 10(1), 11-15, 1991.

McCormack, M. D., Zaucha, D. E., and Dushek, D. W.: First-break refraction event picking and seismic data trace editing using neural networks, Geophysics, 58(1), 67-78, 1993.

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.

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

SEG Wiki: https://wiki.seg.org/wiki/1994_BP_statics_benchmark_model, last access: 22 June 2020.

Sheen, K. L., White, N. J., and Hobbs, R. W.: Estimating mixing rates from seismic images of oceanic structure, Geophysical 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.

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), 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.

Van der Baan, M., and Jutten, C.: Neural networks in geophysical applications, Geophysics, 65(4), 1032-1047, 2000.

Wang, Y.: Frequencies of the Ricker wavelet, Geophysics, 80(2), A31-A37, 2015.

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.~~

~~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.~~

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

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.

Figures

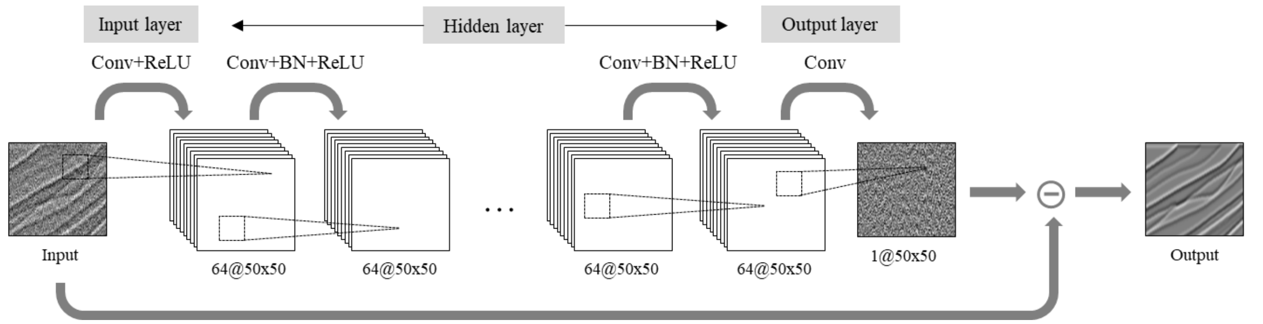


Figure 1: DnCNN architecture. 64@50x50 indicates 64 feature maps with 50x50 size, Conv is two dimensional 3x3 convolution kernel, BN is batch normalization, and ReLU is rectified linear unit activation function.

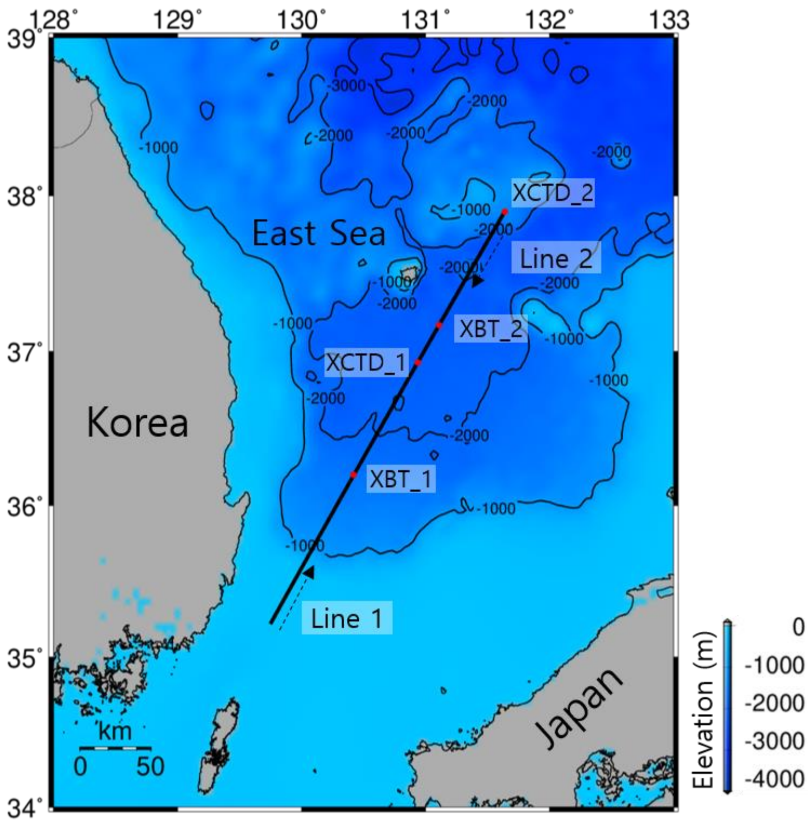
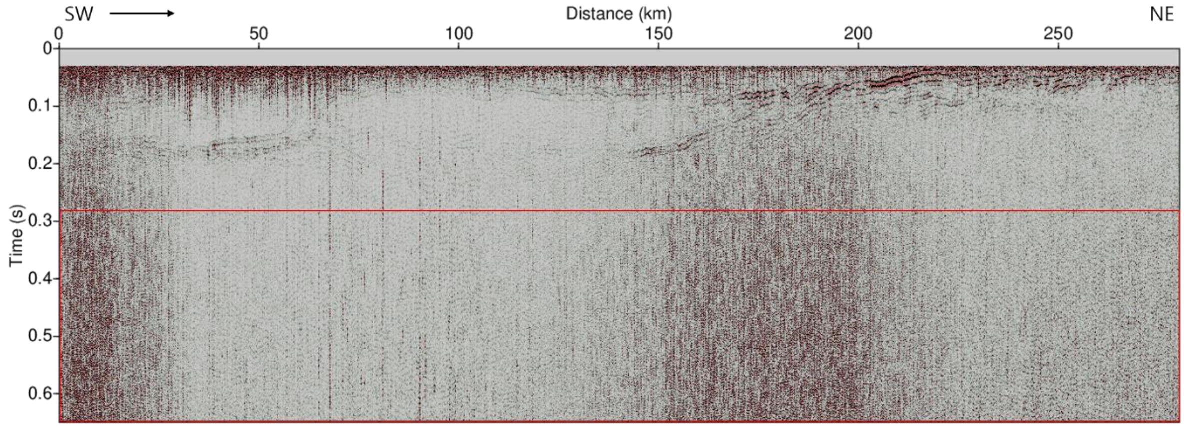
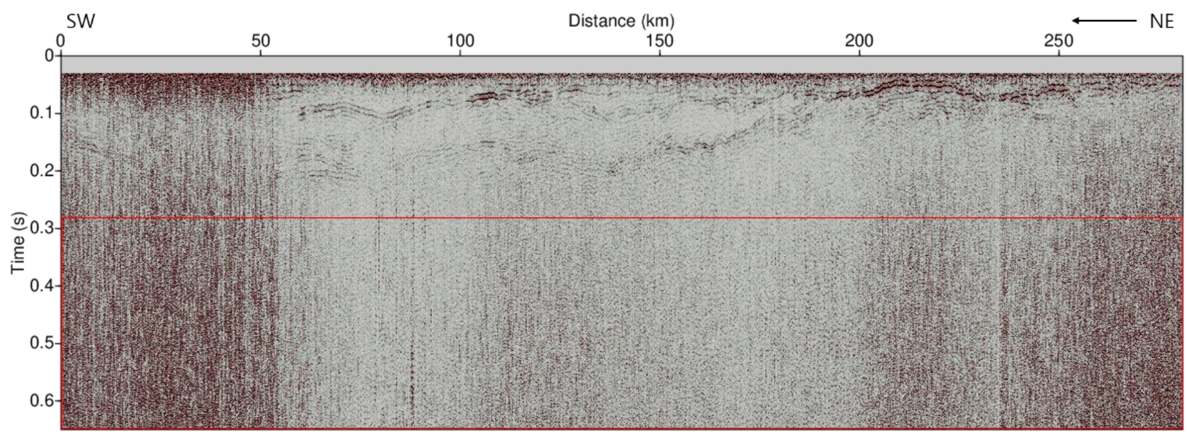


Figure 2: Location of seismic exploration. The black solid line is the survey line, and the black dashed lines with arrow indicate the exploration directions of lines 1 and 2~~.~~, and red dots are the locations of XBTs and XCTDs.

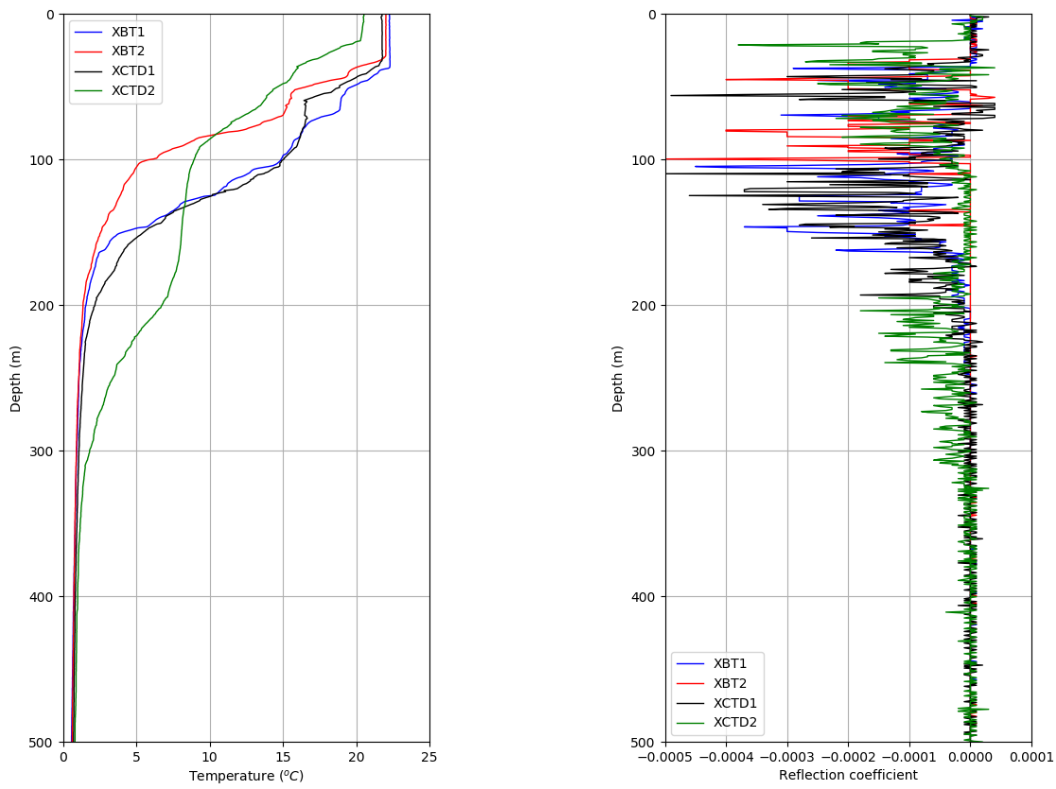


(a)



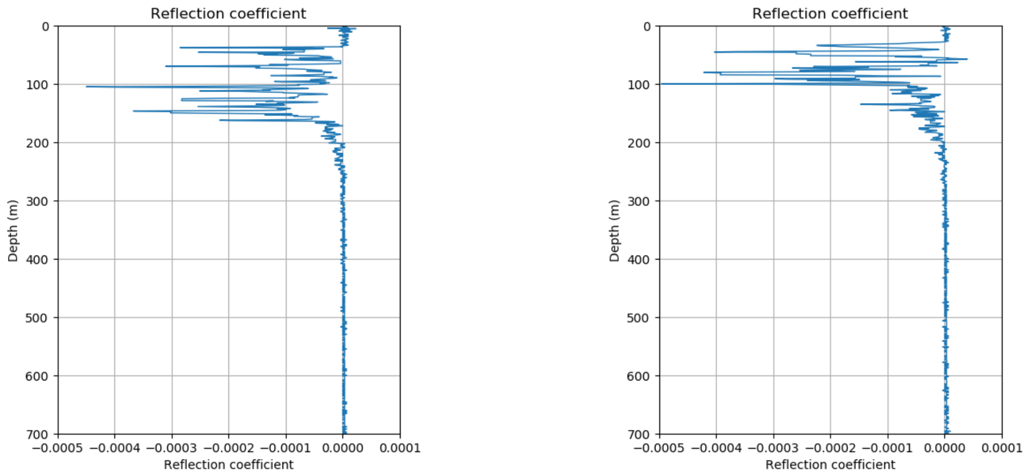
(b)

Figure 3: Processed seismic section of the East Sea: (a) line 1 and (b) line 2. The seismic section in the red rectangle is the noise part used to generate the training data. SW is south west, NE is north east, and black arrow indicates the data acquiring direction.



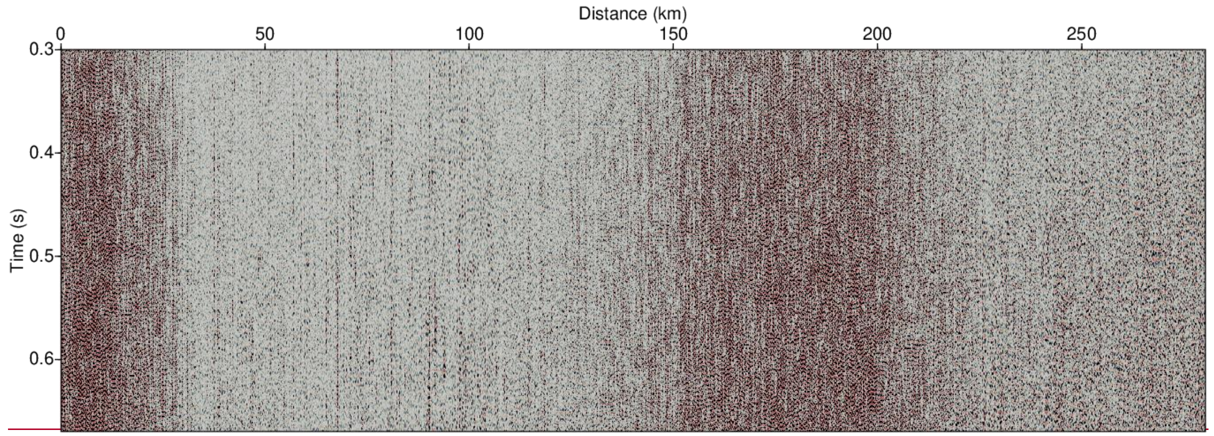
(a) (b)

Figure 4: (a) Temperature and (b) reflection coefficient profiles obtained using 2 XBTs and 2 XCTDs.~~ (a) is XBT_a and (b) is XBT_b in Figure 3.~~

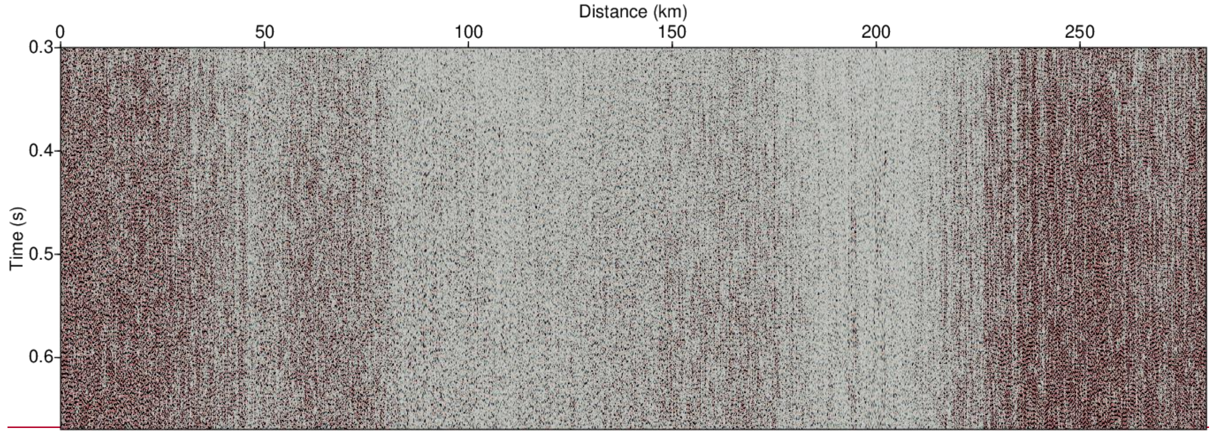


~~(a)~~ ~~(b)~~

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

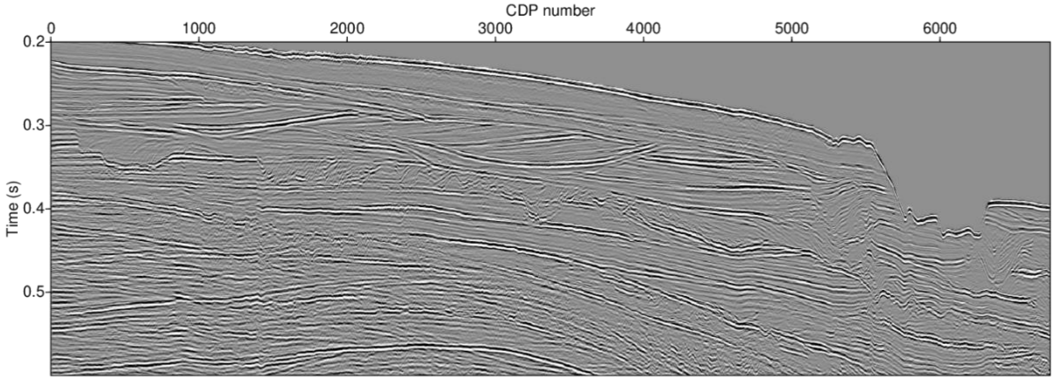


~~(a)~~

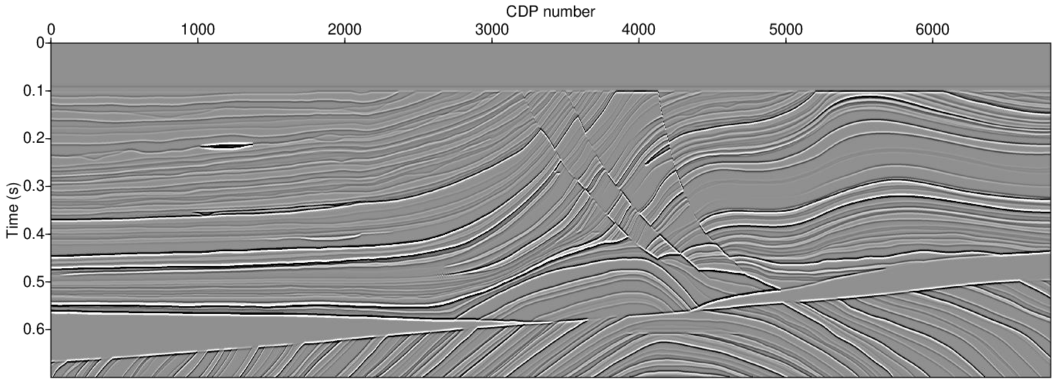


~~(b)~~

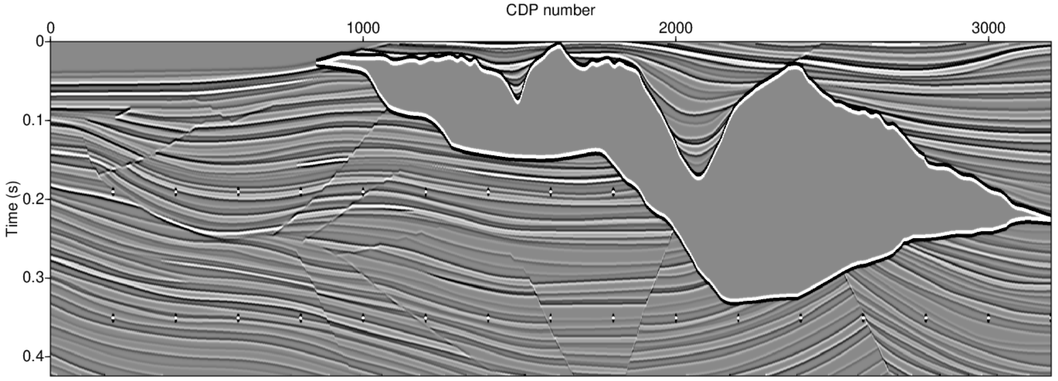
~~Figure 6: Noise section extracted from the East Sea SO data: (a) line 1 and (b) line 2.~~



(a)



(b)



(c)

Figure ~~7~~5: (a) Processed SEZ field seismic section, (b) Marmousi-2 synthetic seismic section and (c) Sigsbee 2A seismic section used to generate the training data.

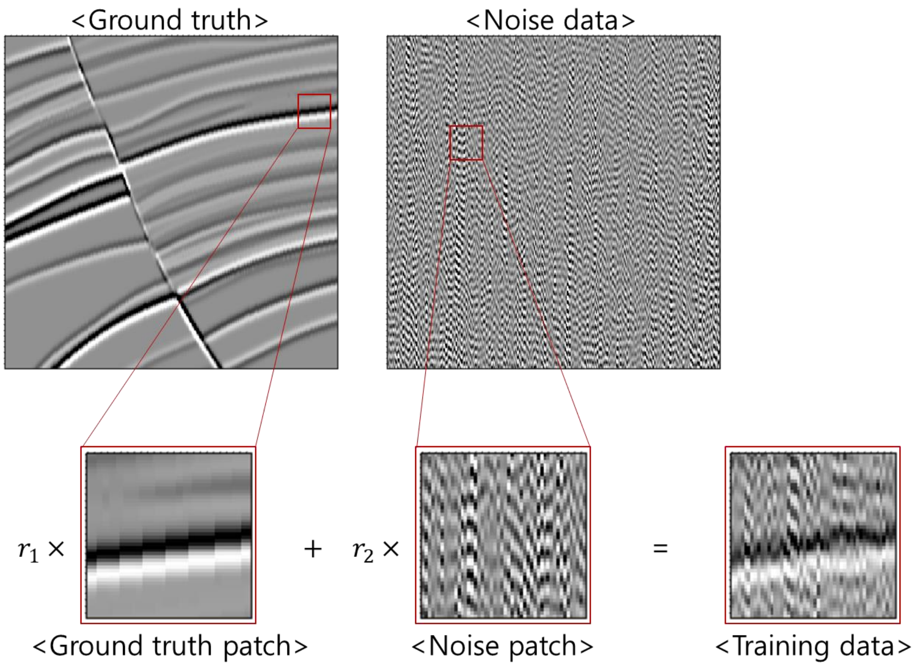


Figure 6~~8~~: Example of constructing the training data.

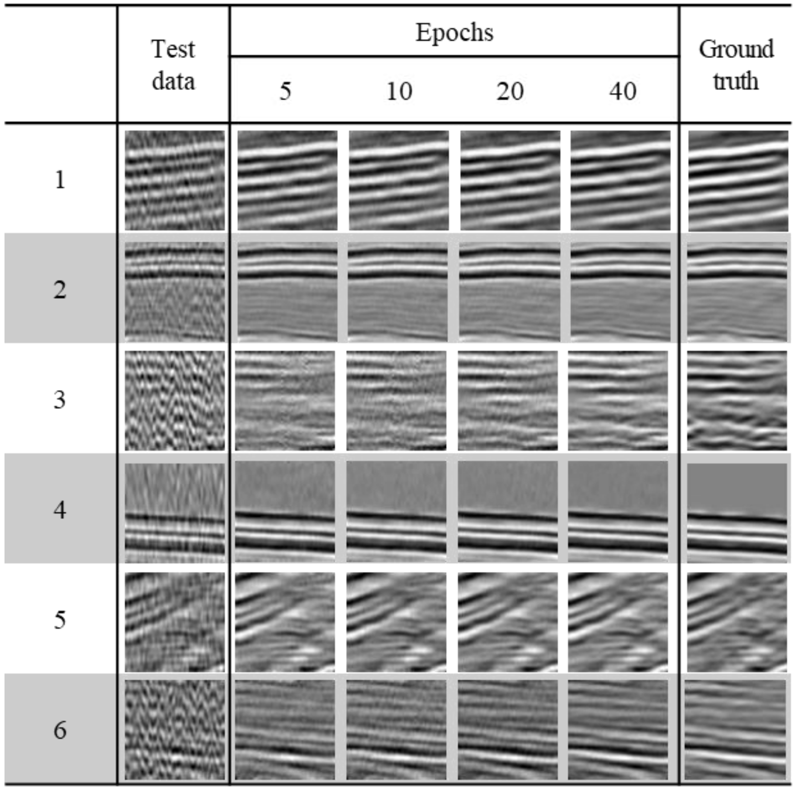
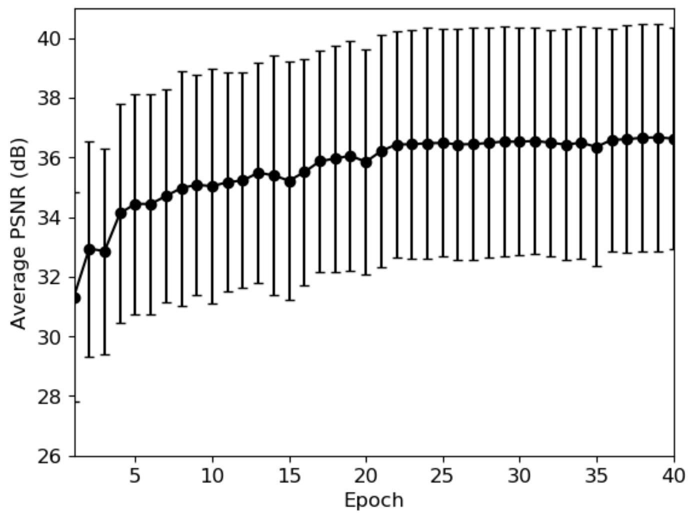
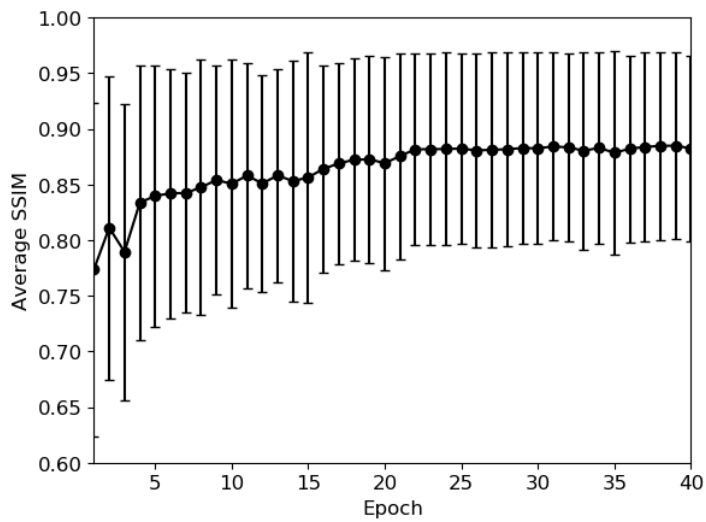


Figure 7~~9~~: Test data, ground truth, and denoised results after applying the DnCNN models trained using training dataset 1.

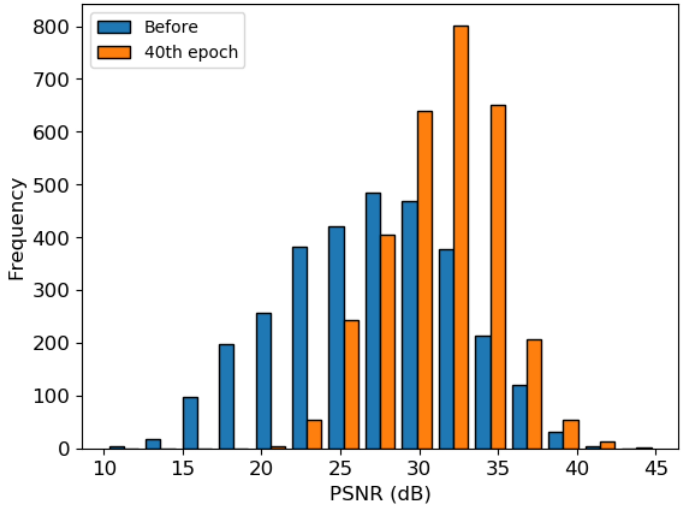


(a)

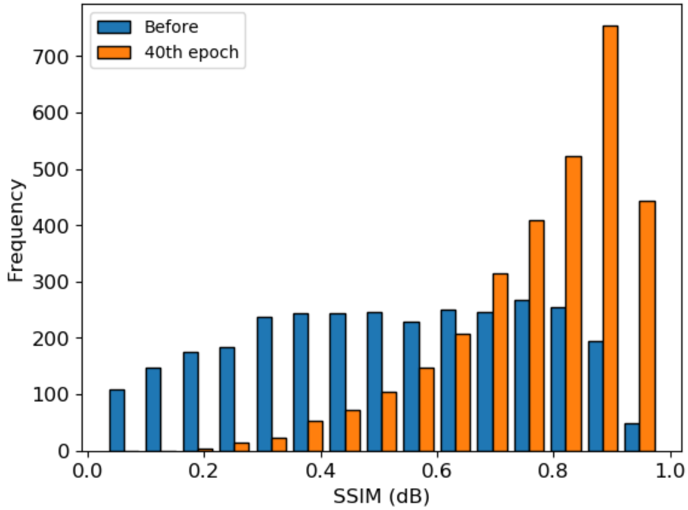


(b)

Figure ~~10~~8: Average (a) PSNR and (b) SSIM with standard deviation of the test result of the first experiment.

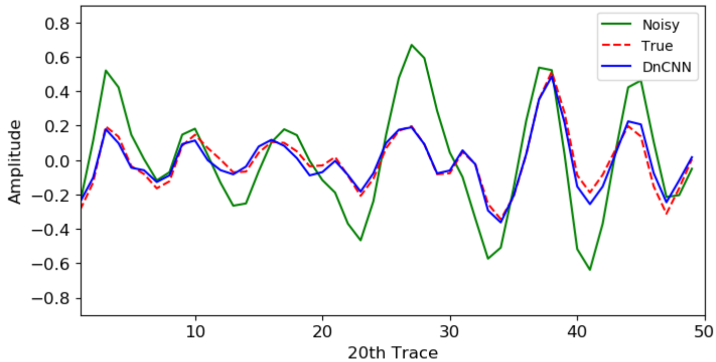


(a)

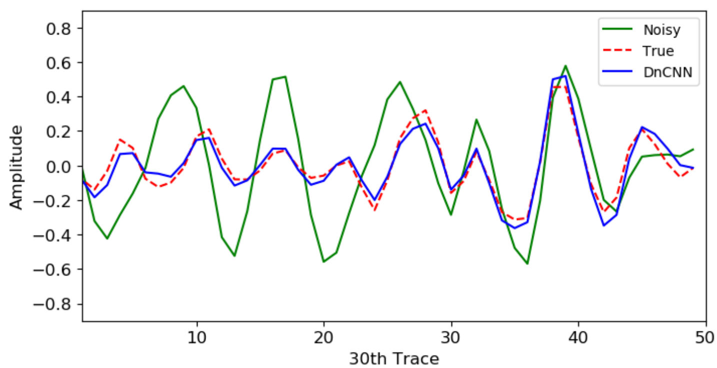


(b)

Figure 9: (a) PSNR and (b) SSIM histogram of the test data before and after applying the 40th epoch of the D1 model.



(a)



(b)

Figure 1~~1~~0: Comparison of the extracted traces before and after applying the 40^{th} epoch of the D1 model.~~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 20^{th} and (b) is the 30^{th} vertical trace of the last test patch in Figure ~~10~~7.

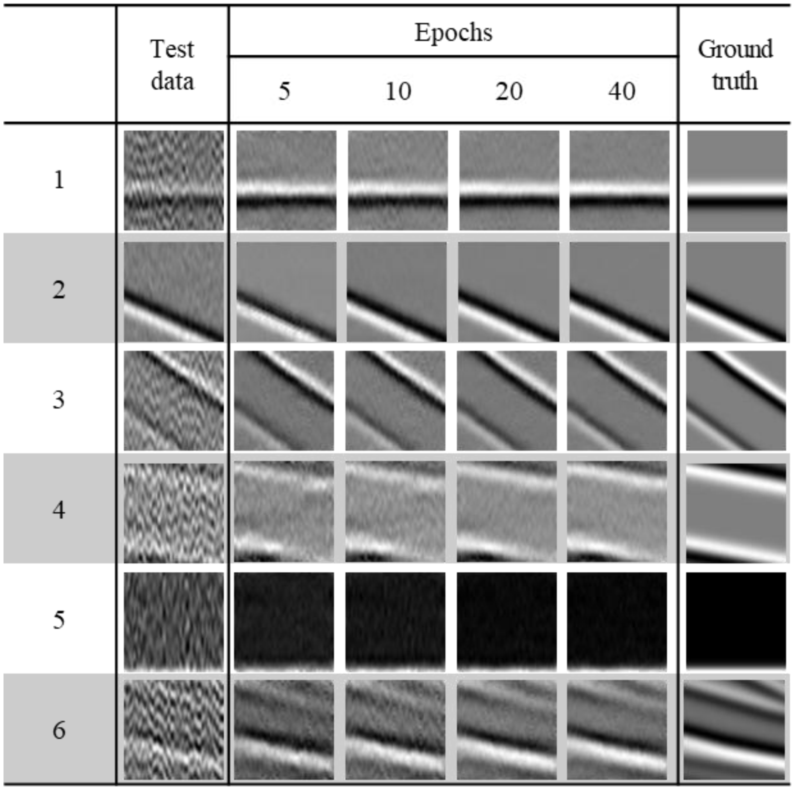
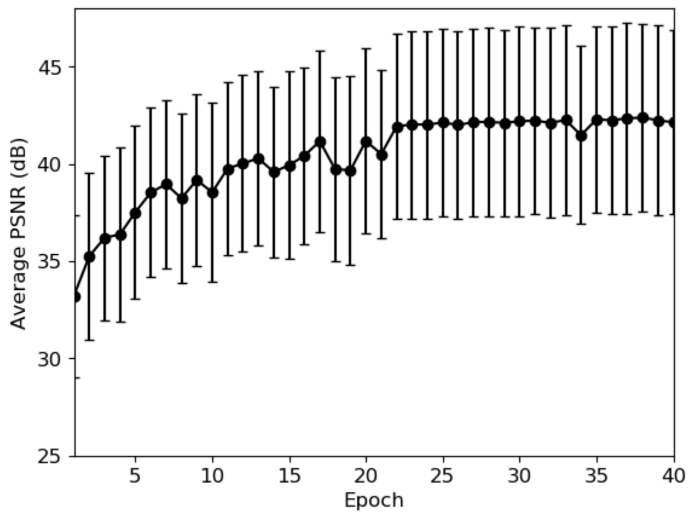
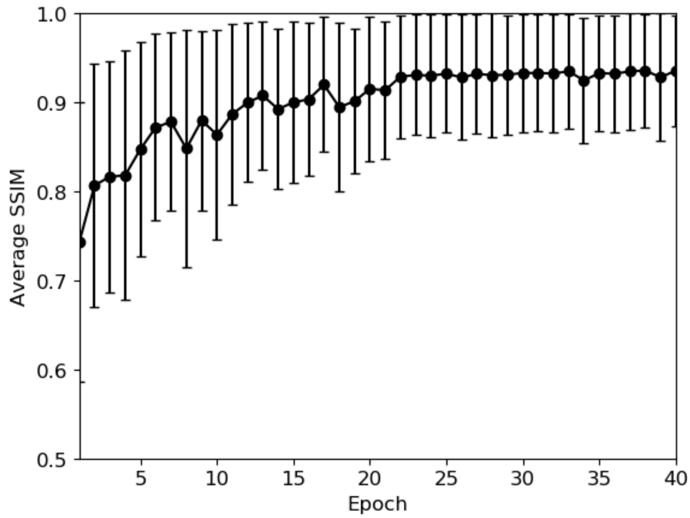


Figure ~~12~~11: Test data, ground truth, and denoised results after applying the DnCNN models trained using training dataset 2.

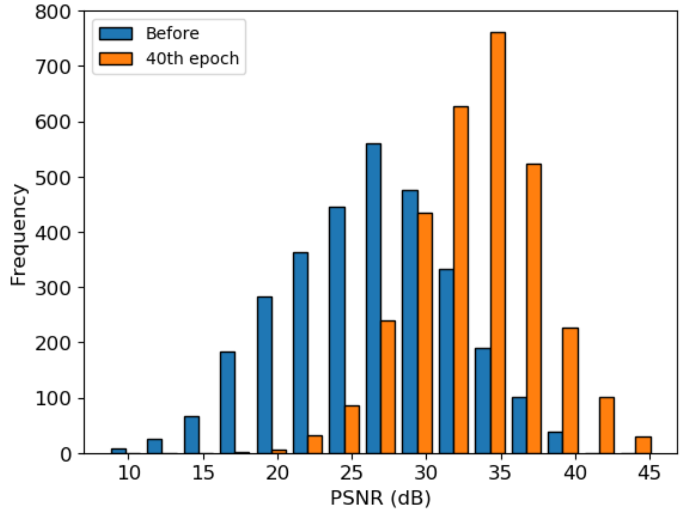


(a)

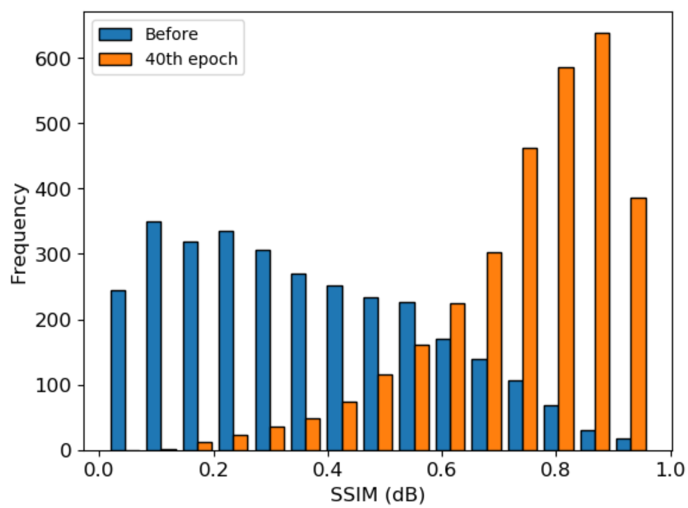


(b)

Figure ~~13~~12: Average (a) PSNR and (b) SSIM with standard deviation of the test result of the second experiment.

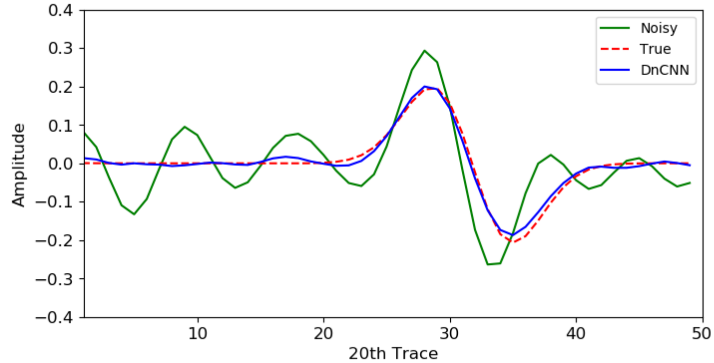


(a)

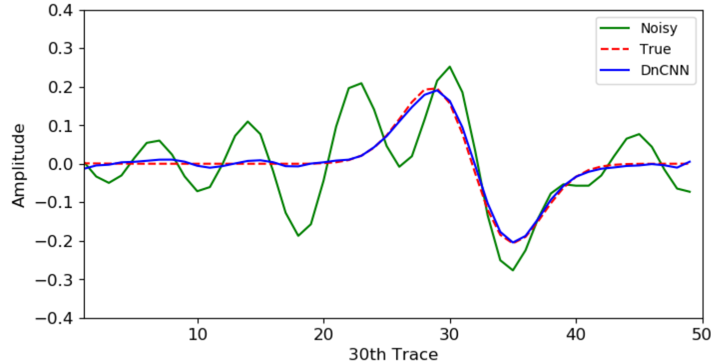


(b)

Figure 13: (a) PSNR and (b) SSIM histogram of the test data before and after applying the 40th epoch of the D2 model.



(a)



(b)

Figure 14: Comparison of the extracted traces before and after applying the 40th epoch of the D2 model.~~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 vertical trace of the first test patch in Figure ~~13~~11.

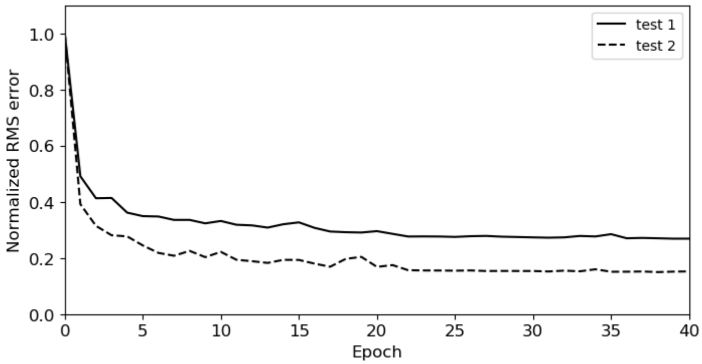


Figure 15: Normalized RMS error between the ground truth and denoised result of the first (solid) and second (dashed) experiments.

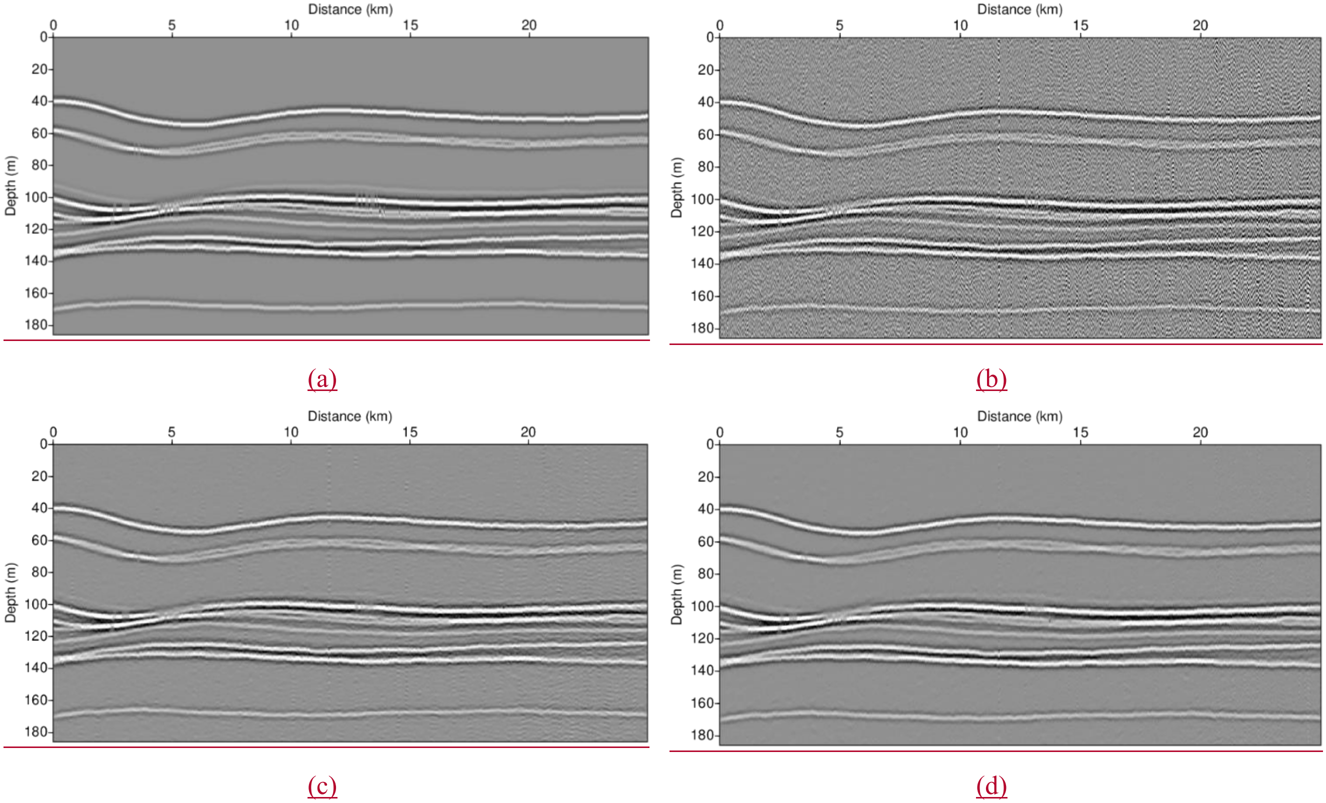


Figure 16: (a) Noise-free and (b) noise-added synthetic water column reflection section and noise-attenuated results using (c) the D1 model and (d) the D2 model.

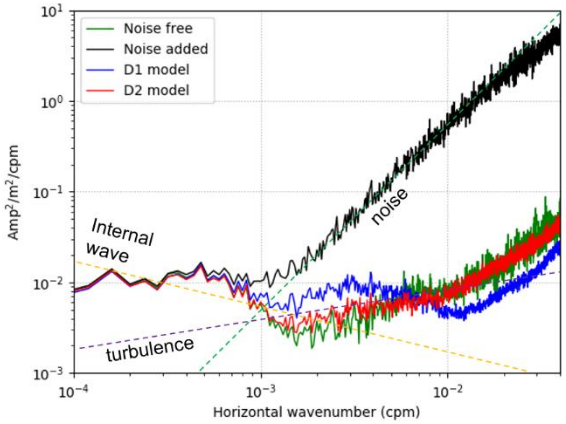
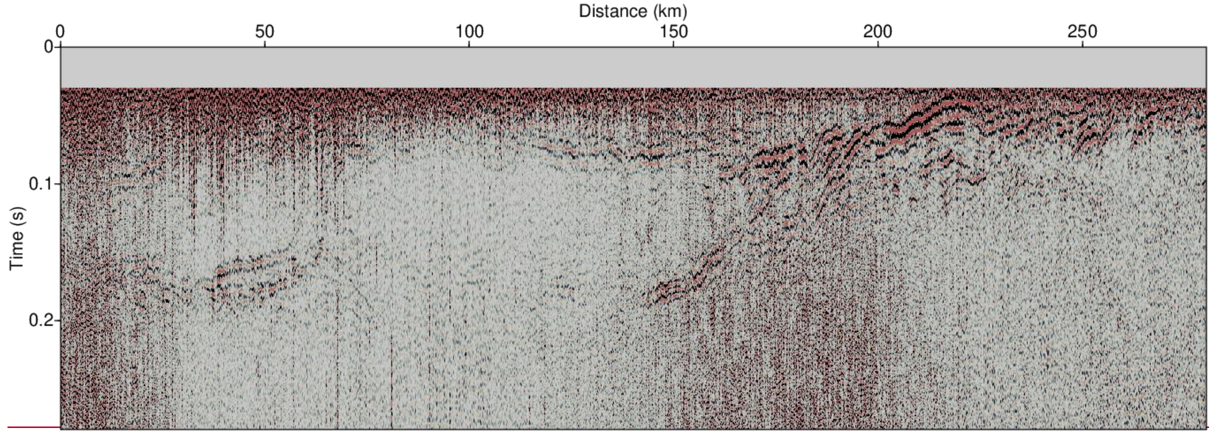
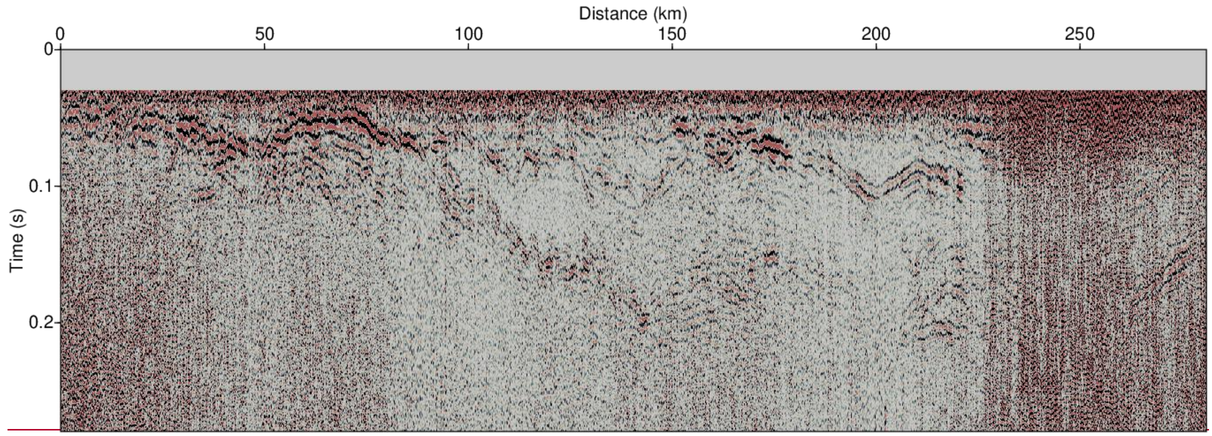


Figure 17: Data slope spectra of noise-free (green) and noise-added (black) synthetic seismic sections and noise-attenuated synthetic seismic section using the D1 model (blue) and D2 model (red).

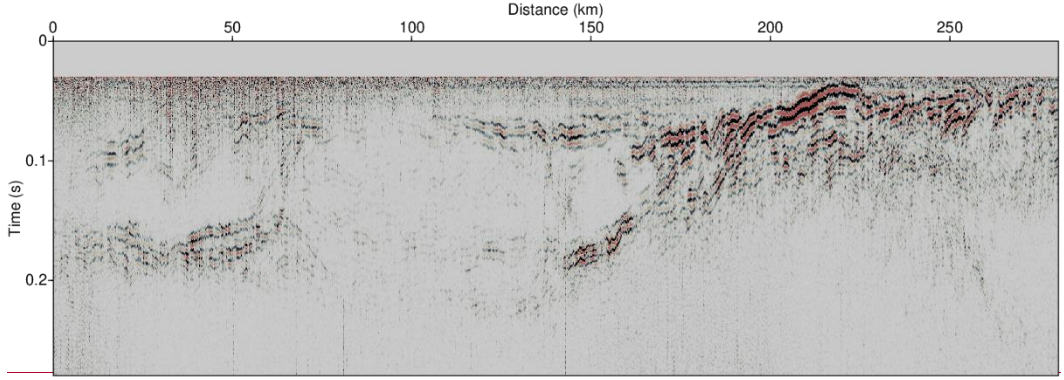


~~(a)~~

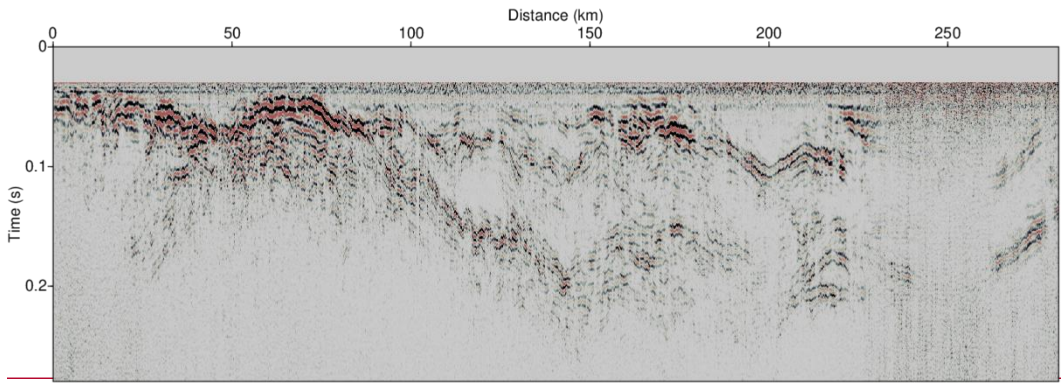


~~(b)~~

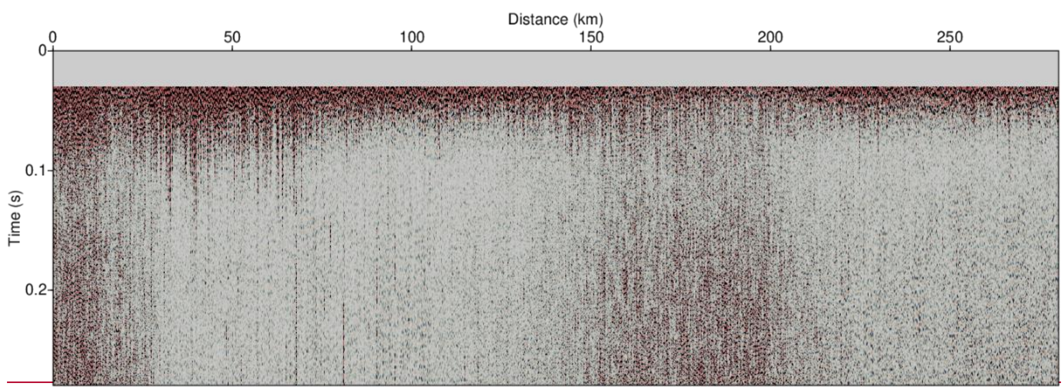
~~Figure 16: (a) Line 1 and (b) line 2 of the East Sea water column seismic section from 0 to 0.28 s.~~


Distance (km)
0
50
100
150
200
250
Time (s)
0
0.1
0.2

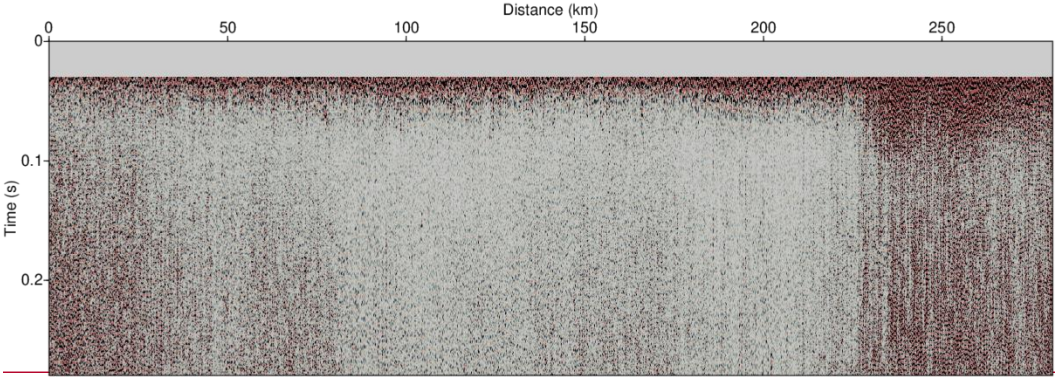
~~(a)~~


Distance (km)
0
50
100
150
200
250
Time (s)
0
0.1
0.2

~~(b)~~

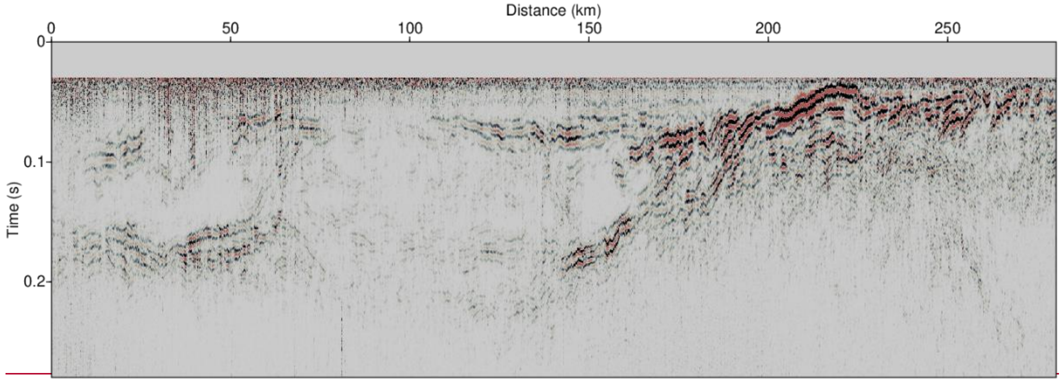

Distance (km)
0
50
100
150
200
250
Time (s)
0
0.1
0.2

~~(c)~~

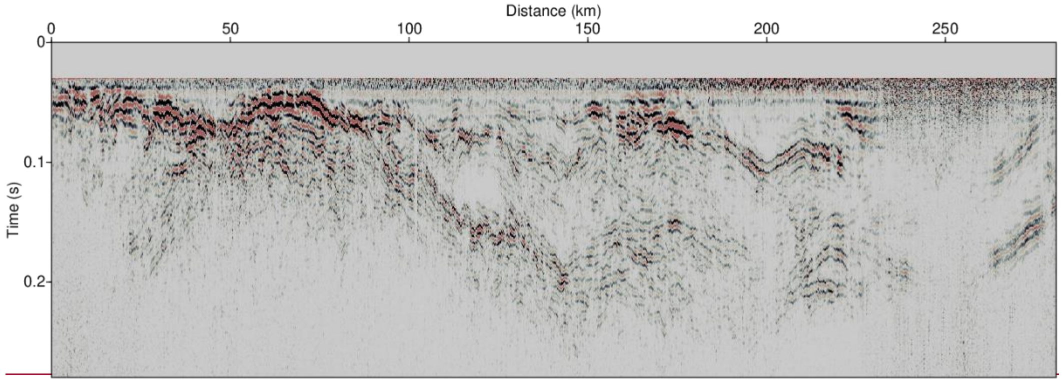


~~(d)~~

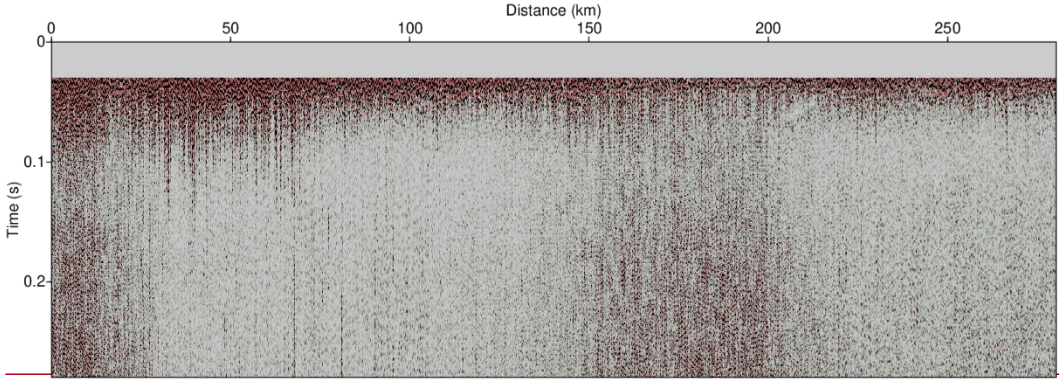
~~Figure 17: Noise attenuation results after applying the D1 model to (a) line 1 and (b) line 2 and estimated noise of (c) line 1 and (d) line 2.~~


Distance (km)
0
50
100
150
200
250
Time (s)
0
0.1
0.2

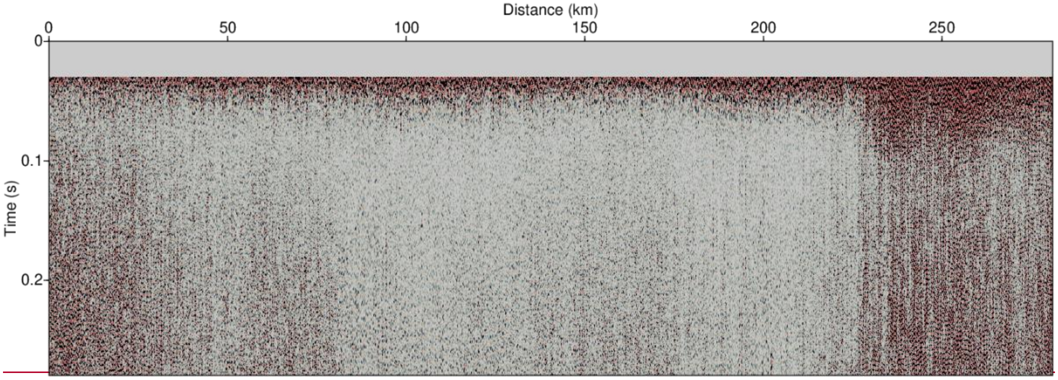
~~(a)~~


Distance (km)
0
50
100
150
200
250
Time (s)
0
0.1
0.2

~~(b)~~


Distance (km)
0
50
100
150
200
250
Time (s)
0
0.1
0.2

~~(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.~~

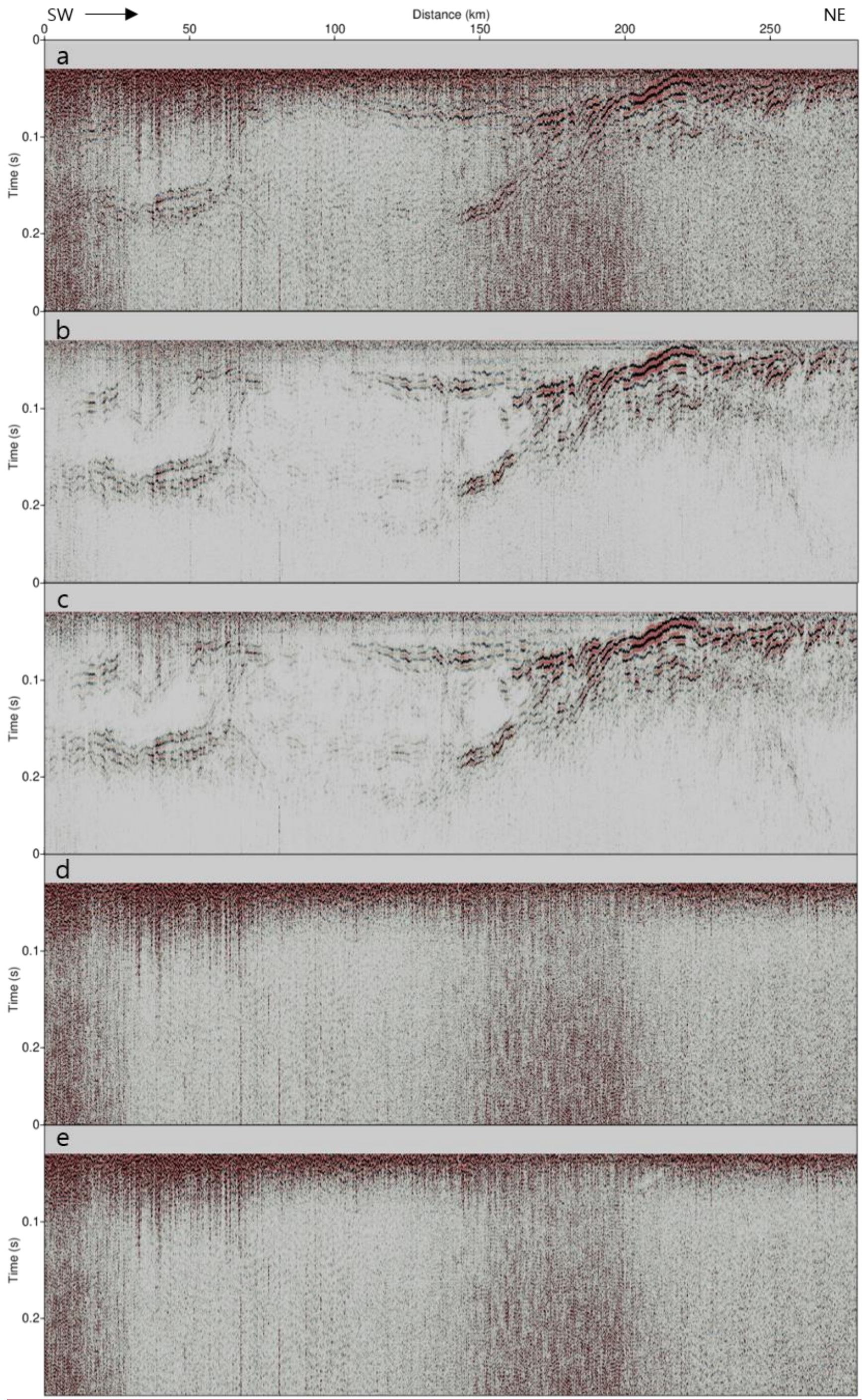


Figure 18: (a) Line 1 seismic section before applying DnCNN, noise-attenuated result using (b) the D1 model and (c) the D2 model, and estimated noise using (d) the D1 model and (e) the D2 model.

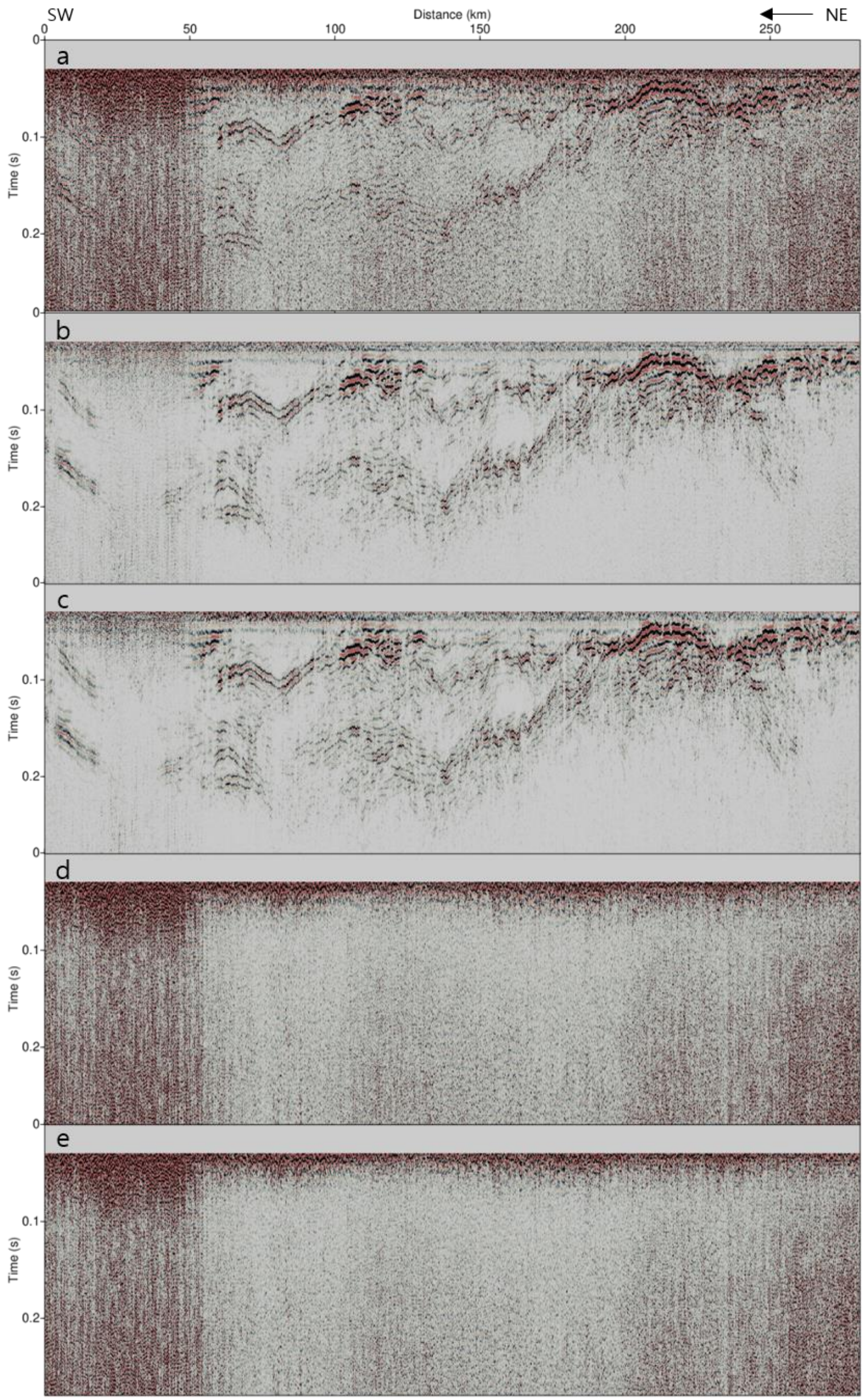
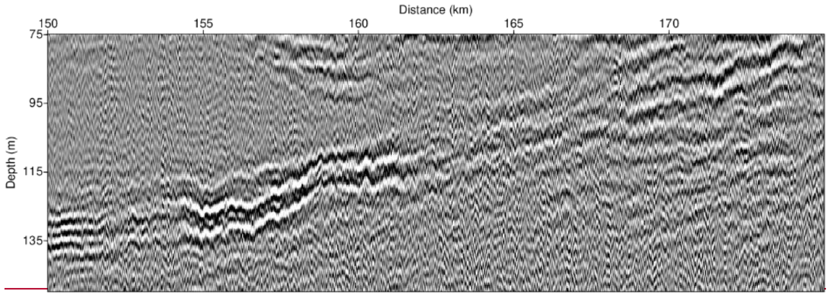
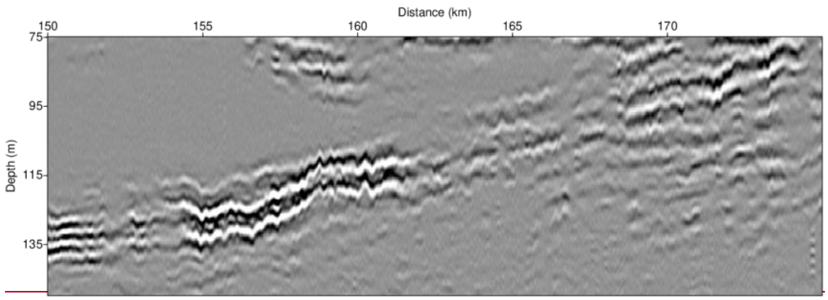


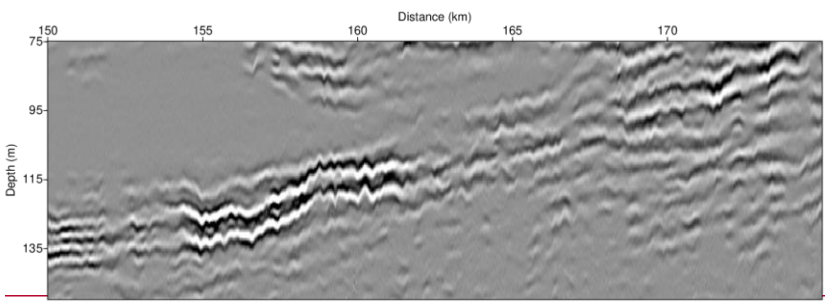
Figure 19: (a) Line 2 seismic section before applying DnCNN, noise-attenuated result using (b) the D1 model and (c) the D2 model, and estimated noise using (d) the D1 model and (e) the D2 model.


Distance (km)
150
155
160
165
170
75
95
115
135
Depth (m)

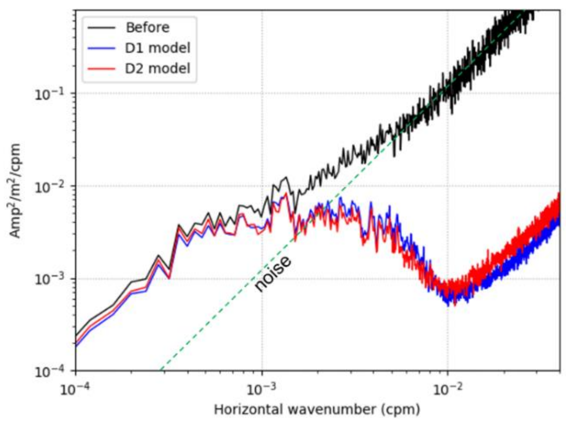
~~(a)~~


Distance (km)
150
155
160
165
170
75
95
115
135
Depth (m)

~~(b)~~

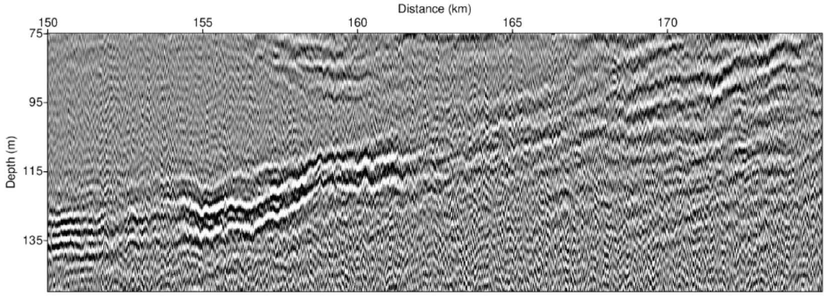

Distance (km)
150
155
160
165
170
75
95
115
135
Depth (m)

~~(c)~~

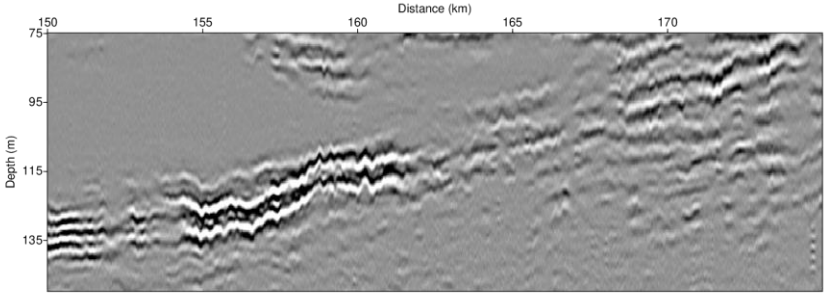


~~(d)~~

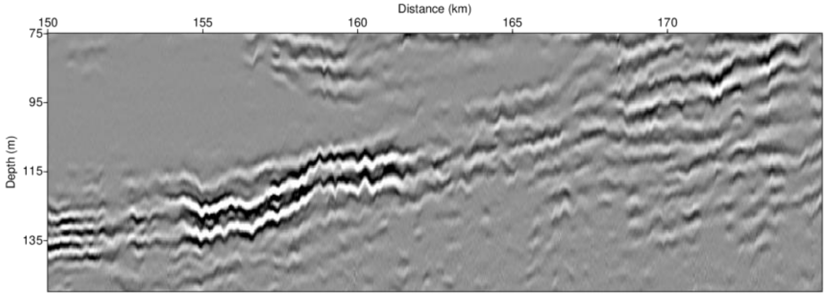
~~Figure 19: Extracted seismic sections ((a) is the section before noise attenuation, (b) is the section after applying the D1 model and (c) is the section after applying the D2 model) and (d) shows calculated data slope spectra of (a), (b) and (c).~~


Distance (km)
150
155
160
165
170
75
95
115
135
Depth (m)

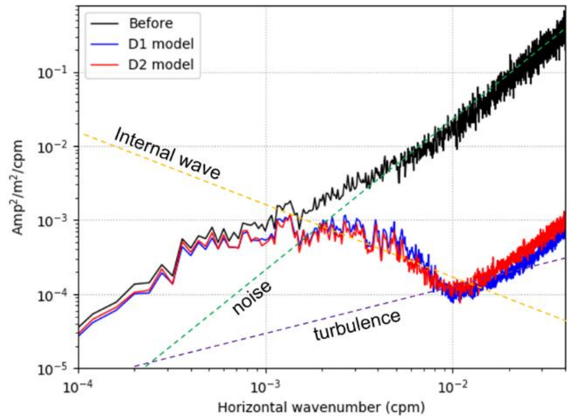
(a)


Distance (km)
150
155
160
165
170
75
95
115
135
Depth (m)

(b)


Distance (km)
150
155
160
165
170
75
95
115
135
Depth (m)

(c)



(d)

Figure 20: 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). (d) shows the calculated data slope spectra of (a), (b) and (c).