Shear wave noise attenuation in ocean bottom node data using machine learning

T. Seher[1], G. Yalcin[2], M. Roberts[1], Y. Ren[1]

[1] TGS; [2] University of Leeds

Summary

Noise is an ongoing issue in seismic processing and our capability to successfully process ocean bottom node data relies on our capability to successfully attenuate the presence of shear wave noise in vertical geophone recordings. To attenuate this noise, we have previously relied on co-denoise techniques in various transform domains. However, these methods can be hard to parametrize and costly to apply in practice. In this study, we demonstrate that machine learning (ML) algorithms can be used for shear wave noise attenuation at least as a fast-track solution. This study presents two new findings. First, we show that for shear wave noise attenuation, ML solutions using only the vertical geophone perform as well as dual-component solutions using both the hydrophone and geophone. Second, we analyse the generalizability of ML solutions. Trained ML networks are shown to generalize to seismic data from a different (unseen) seismic experiment not included during training.

Shear wave noise attenuation in ocean bottom node data using machine learning

Introduction

Noise is an ongoing issue in seismic processing of ocean bottom node (OBN) data, and shear wave contamination of the compressional wave records can still pose a processing challenge. A variety of processing techniques have been developed to attenuate this noise using thresholding in various transform domains (Craft & Paffenholz, 2007; Yang et al., 2020; Ren et al., 2022). In recent years, machine learning (ML) techniques have become popular for seismic denoise such as swell noise attenuation (Brusova et al., 2021; Valenciano et al, 2022). Similarly, ML can be successfully applied to attenuate shear wave noise (i.e., Vz noise) in OBN records (Sun et al., 2023; Yalcin, 2023).

In this study, we present three findings. First, we demonstrate that ML denoise can be applied in a fast-track processing sequence to attenuate Vz noise. Replacing conventional solutions with an ML solution is attractive because of the potential to reduce human effort. Second, we demonstrate that there is no advantage in using both the hydrophone and geophone component during the training of our ML denoise networks as compared to using the geophone component only. This result is contrary to what we expected from conventional seismic processing where the hydrophone component is crucial for successful co-denoise. Third, we demonstrate that given sufficient training data, ML networks can make accurate noise predictions for seismic experiments not included in the training data (i.e., the networks are able to generalize).

In this study, we first review the curvelet Vz denoise approach applied to generate the training data and describe the ML network architecture. Next, we compare the network performance using single and dual component training data and demonstrate that ML denoise can be part of a fast-track processing sequence. We then apply our ML denoise network to seismic data from an experiment not included in the training set to test its generalizability. We end this paper with a few observations.

Method

The vertical geophone data is usually dominated by strong shear wave noise whereas the hydrophone data is mostly free of this noise. Most shear wave noise attenuation algorithms exploit this property to attenuate the shear wave noise by applying a co-denoise procedure using the hydrophone as a guide in denoising the geophone. In this study, we chose a curvelet domain technique. The output of this approach is then used as the training data for our ML denoise process.

The curvelet domain is a sparse representation of the data into fine scales and angles where the noise can be easily distinguished from the signal. We used the method of Vz denoise in the curvelet domain proposed by Ren et al. (2022), in which the hydrophone component is used as a guide in the denoising process. Both the hydrophone and vertical geophone components of the data are first transformed into the 2D curvelet domain. For each scale and orientation panel in the curvelet domain, the P/Z amplitude ratios are compared to an adaptive thresholding value targeting the Vz noise. When the P/Z amplitude ratio is lower than the threshold, the Vz coefficient is set to zero, while for higher ratios the Vz coefficients are unchanged. The inverse transform is then applied to obtain the denoised Vz component.

The ML model used in this study was a deep convolutional U-net, with skip connections and a variational autoencoder for the bottleneck layer (Brusova et al., 2021). Differing from Brusova et al. (2021), data was trained with random cropping and flip augmentation to prevent overfitting to the training dataset and to improve generalizability to unseen data. The model was trained with the Adam optimizer and a step LR scheduler. For tests with both hydrophone and vertical component input, the two components were input as separate input channels in the network.

To create training data for the ML network, we have selected data from two recent experiments in the North Sea (Sleipner & NOAKA). From both experiments we selected a subset of 12 receivers with a full 3D shot carpet. We then applied conventional 2D curvelet denoise to all 24 receiver gathers treating each shot line separately. The denoise procedure outputs both a signal estimate and a corresponding noise model, which we use as part of our ML training procedure. We then apply the network for noise attenuation of a third experiment in the North Sea (Utsira) not included in training.

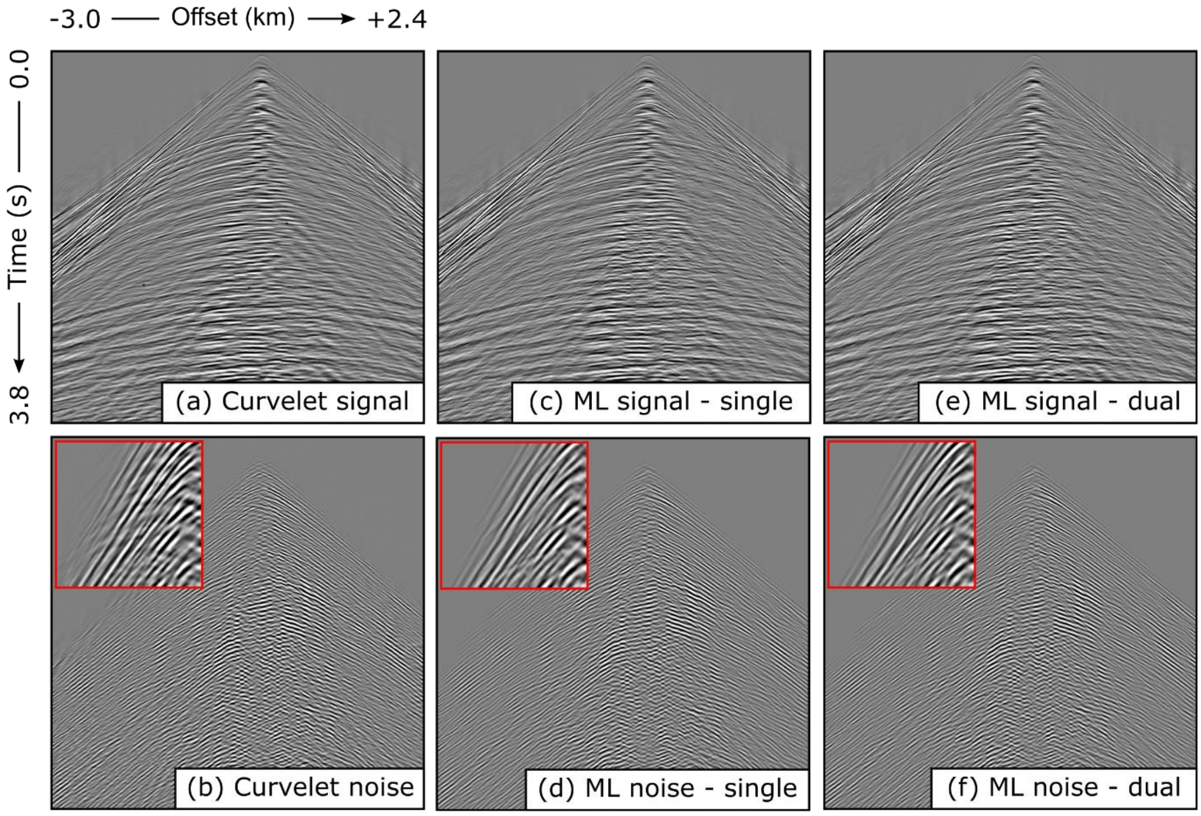

Figure 1 *Receiver gather from the Sleipner experiment comparing the signal and noise model extracted using 2D curvelet Vz denoise (a&b), our single component ML Vz denoise solution (c&d) and the dual component ML Vz denoise solution (e&f). The inset highlights curvelet transform artefacts.*

Single or Dual Component Training Data

One of the objectives of this study was to understand whether there was an advantage to using both hydrophone and geophone components as input to the ML denoise training process. To this end, we trained two networks. The single component network used only the geophone records as input and the noise models from the 2D curvelet denoise step as labels. The dual component network used both the hydrophone and geophone as input and the noise models as labels.

Both the single and dual component networks successfully learnt to attenuate Vz noise in vertical geophone records (Figure 1). Comparing the ML results to the curvelet results for a receiver not included in the training data allows two interesting insights. First, the ML results appear cleaner; the ML network has not learnt the transform artefacts. Second, the ML results appear more conservative; the amplitudes of the noise model are slightly lower, and some events are not identified as noise. Comparing the single and dual component results (Figure 1) shows only small differences. Contrary to our expectation, there is no advantage in using multi-component data as part of the ML based denoise procedure. All required information appears contained in the input geophone data and the noise labels.

To illustrate the performance of the curvelet and ML denoise approaches, we compare their impact on a receiver stack for the Sleipner experiment (Figure 2). Both approaches successfully mitigate shear wave noise, and no primary damage is obvious. While the ML denoise approach has left behind some steeply dipping events, the resulting image appears “crisper” than the curvelet solution. While this confirms that ML solutions currently struggle to match the quality of the training data, the noise attenuation appears sufficient for fast-track processing.

Generalization Ability

Can we apply existing networks to new experiments not part of the training data without the need for extensive retraining? To answer this question, we applied our ML denoise network to a seismic receiver from the Utsira region of the North Sea. We stress that no data from the Utsira region was included in

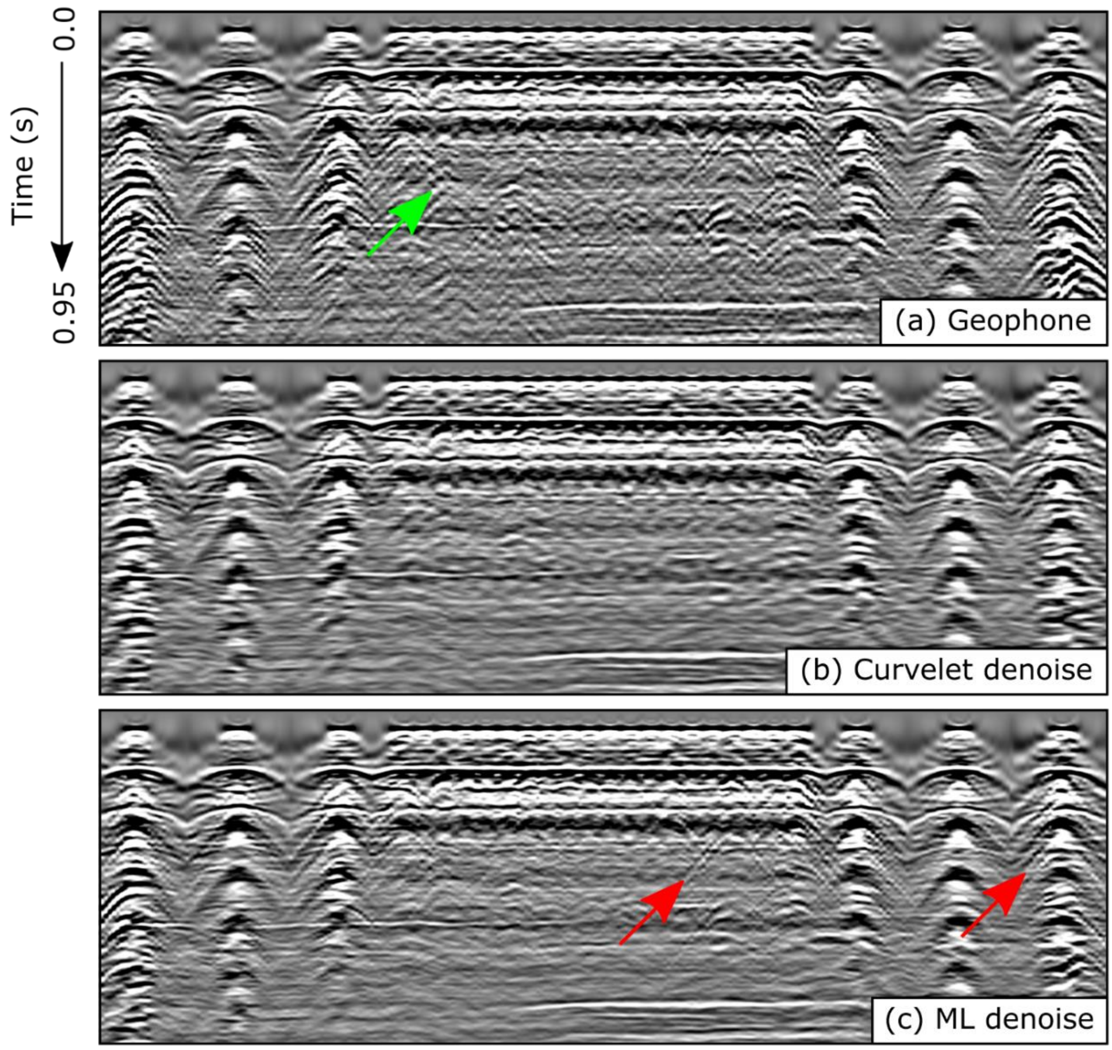

Figure 2 *Receiver stacks of seismic data from the Sleipner OBN experiment. Both the 2D curvelet Vz denoise algorithm (b) and the ML Vz denoise solution (c) attenuate the shear wave noise present in the raw vertical geophone data (green arrow in a). The red arrows highlight locations where the curvelet solution outperformed the ML solution.*

training the network and different node types were used during the acquisition. Comparing the ML Vz denoise results to conventional Vz denoise using the 2D curvelet transform (Figure 3) demonstrates that the pre-trained ML network successfully attenuates a large part of the Vz noise apparent in the geophone data. Crucially, the ML denoise performance may be improved through additional training data. This has the interesting implication that quality and quantity of training data rather than the algorithm might determine the success of ML based seismic processing.

Conclusions

In this study, we have presented an ML based approach for shear wave noise attenuation in OBN data. We have demonstrated that the approach delivers results that are sufficient for a fast-track processing sequence. Interestingly, the network does not appear to learn the transform artefacts from the training data. In addition, our denoise results have always been more conservative than the approach used in creating the training data. Most importantly, this study has allowed us two new insights. First, contrary to our expectations, there is no inherent advantage to using multi-component input data. Second, our networks can predict shear wave noise for experiments that were not part of the training data. This study is a step towards applying ML based denoise techniques as part of standard processing sequences and reducing the human effort. Crucially, ML based solutions are easier to apply, and parametrization of the algorithms is only required once during the creation of the training data.

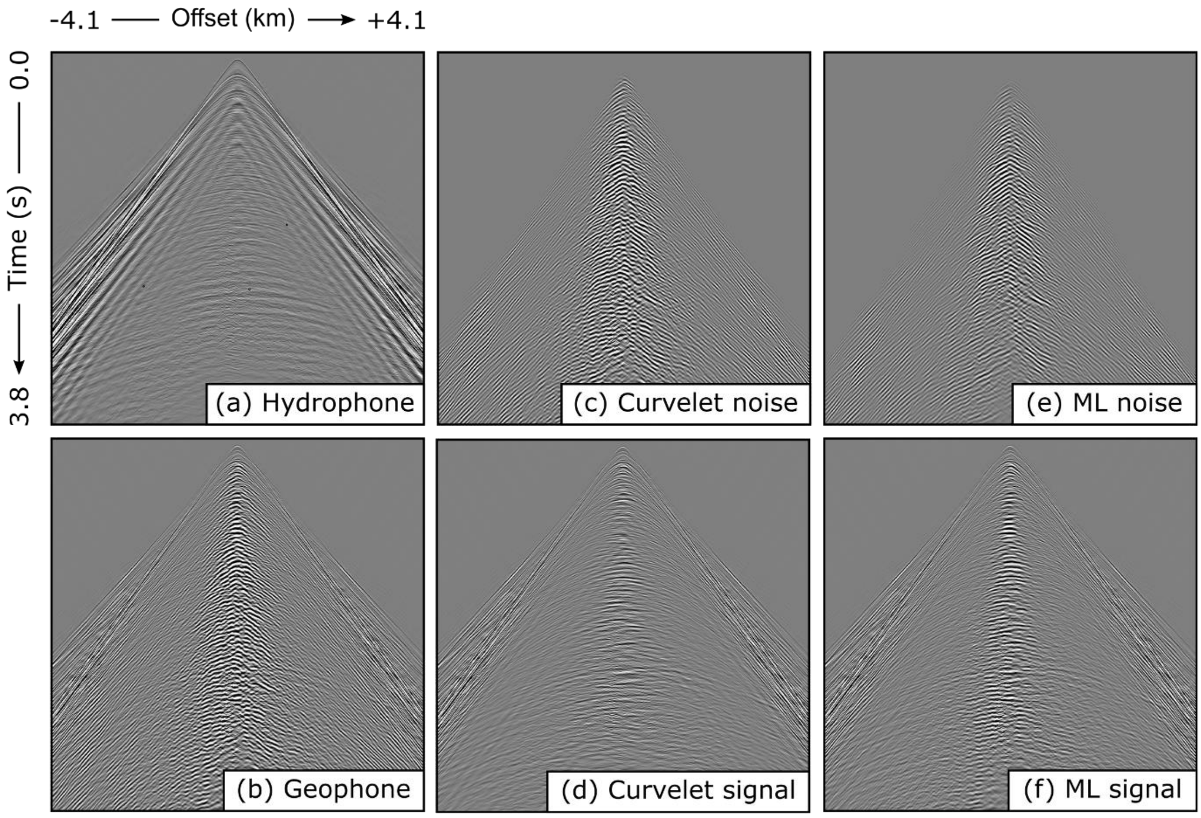

Figure 3 *Receiver gather from the Utsira experiment not included in the training set. The input hydrophone and geophone data (a&b) are used in 2D curvelet Vz noise attenuation (c&d). The ML based noise attenuation (e&f) uses only the geophone data during inference and compares well with the curvelet results (c&d).*

Acknowledgements

This work was supported by many TGS colleagues, most importantly O. Brusova, E. Cho, G. Szumski, C. Ursenbach, and Z. Liu. We are thankful to TGS management and our joint venture partners (CGG and Axxis Multi Client, an Aquila Holdings company) for permission to show the data.

References

Brusova, O., Poche, S., Kainkaryam, S., Valenciano, A. and Sharma, A. [2021] An innovative strategy for seismic swell noise removal using deep neural networks. *First International Meeting for Applied Geoscience & Energy*, Expanded Abstracts, 3179–3183. https://doi.org/10.1190/segam2021-3592770.1

Craft, K.L. and Paffenholz, J. [2007] Geophone noise attenuation and wavefield separation using a multidimensional decomposition technique. *SEG Technical Program Expanded Abstracts*, 2630–2634. https://doi.org/10.1190/1.2793013

Ren, Y., Yang, C., Seher, T., Hawke, M. and Cho, E [2022] Vz denoise and P/Z matching in 3D curvelet domain. *Second International Meeting for Applied Geoscience & Energy*, Expanded Abstracts, 2862–2866. https://doi.org/10.1190/image2022-3745669.1

Sun, J., Jafargandomi, A. and Holden, J. [2023] Deep learning-based Vz-noise attenuation for OBS data. *Third International Meeting for Applied Geoscience & Energy*, Expanded Abstracts, 1480–1484. https://doi.org/10.1190/image2023-3910189.1

Valeciano, A., Brusova, O. and Cheng, C. [2022] Efficient Swell Noise Removal Using a Global Deep Neural Network Model. *First Break*, **40**, 51–55. https://doi.org/10.3997/1365-2397.fb2022012

Yalcin, G.B. [2023] *Vz Noise Attenuation Using Machine Learning for OBN Data*. MSc thesis, Leeds University.

Yang, C., Huang, Y., Liu, Z., Sheng, J. and Camarda, E. [2020] Shear wave noise attenuation and wavefield separation in curvelet domain. *SEG Technical Program Expanded Abstracts*, 1805–1809. https://doi.org/10.1190/segam2020-3426418.1