Enhancing Seismic Data Quality: A Machine Learning Approach to Denoising and Signal Damage Reduction

Mark Roberts, Olga Brusova, Leandro Gabioli, and Alejandro Valenciano, TGS*

Summary

We propose a denoise workflow comprising a supervised ML (Machine Learning) model applied in the common shot domain and a self-supervised ML signal-add back model in the common channel domain. The supervised ML-based denoise (Brusova *et al.*, 2021) learns from training data containing recorded noise and provides robust, high-quality results that can successfully tackle various noise types in a single pass without requiring complex parameterization. However, some signal leakage can occur, producing primary damage. We demonstrate a self-supervised signal add-back technique based on the blind-trace network (Birnie *et al.*, 2021) that mitigates the primary damage and produces a complete denoise solution. The technique is applied to various vintages of offshore streamer data. The results from our novel workflow show significant improvements in recovery/denoise of low frequency (<3 Hz) on legacy streamer data, which is critical for success in deghosting/broadband processing and FWI.

Introduction

Seismic data from offshore marine streamer data can contain many types of background noise, including swell noise caused by waves and swell on the ocean surface, tug noise caused by the cable being jerked by the tow vessel, propeller noise from the streamer vessel or other nearby ships, cable noise from mechanical vibrations along the cable, etc. Traditional denoise techniques have involved multiple passes of time-filtering algorithms in multiple domains to try randomizing and removing the noise, such as time-frequency vector median filtering (Seher and Ortega, 2018). While such approaches can be effective, they typically require significant effort to optimize parameters, and a given set of parameters only provides acceptable results over a narrow range of noise conditions.

Brusova *et al.* (2021) demonstrated the potential of using Convolutional Neural Networks (CNNs) to accurately estimate swell noise for streamer data. Additional field noise records have been added in this work to increase the generalizability to more noise types including linear noise such as cable-tug and paravane noise.

Several papers have been written on self-supervised image denoising including noise2noise (Lehtinen *et al.*, 2018) where input noisy data had additional noise applied and noise2void (Krull *et al.*, 2019) where a network was designed to recover masked input pixels. An improvement in the noise2void implementation came from Laine *et al.* (2019), who developed a network with a convolutional operator shifted to only have a receptive field to one side of the output pixel. By applying the model with shifts applied four ways (e.g., up, down, left, and right), they ended with a model that had a receptive field that, for a single pass, is blind only to the corresponding input pixel, referred to as the Blind Spot network (BSN). This is very efficient and can use every pixel in an image for training. The noise2void model works under the assumption that noise signals are pixel-wise independent and have a zero mean.

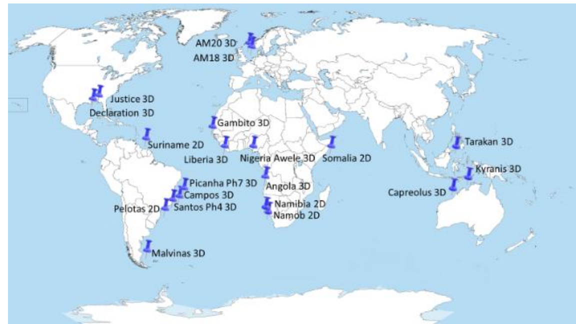

Figure 1: Location of data used during model training.

Seismic noise is frequency-limited and therefore rarely pixel-wise independent, so the BSN is not applicable. However, in the correct domain, noise is often trace-wise independent. An example of this is tug noise, which is coherent in the common-shot domain but incoherent in the common-offset or common-receiver domain. While this blind-trace network (BTN) approach is highly successful at attenuating noise, it was observed to have significantly higher signal damage than supervised-based approaches that estimate noise.

Birnie *et al.* (2021) implement a blind-trace self-supervised denoise flow for seismic data based on an extension of the masked base noise2void paper (Krull *et al.*, 2019). Luiken *et al.* (2023) use a full blind-trace network based on Laine *et al.* (2019), as a component of an inversion-based deblending.

The key idea presented here is to utilize the blind trace network to find residual coherent signal in the estimated noise data, rather than to estimate the noise directly. To accomplish this, the predicted supervised-ML noise records for a sail line are sorted into common-channel domain. The blind trace network is next trained on the noise records and learns to predict the residual coherent signal (leaked signal).

A Machine Learning Approach to Denoising and Signal Damage Reduction

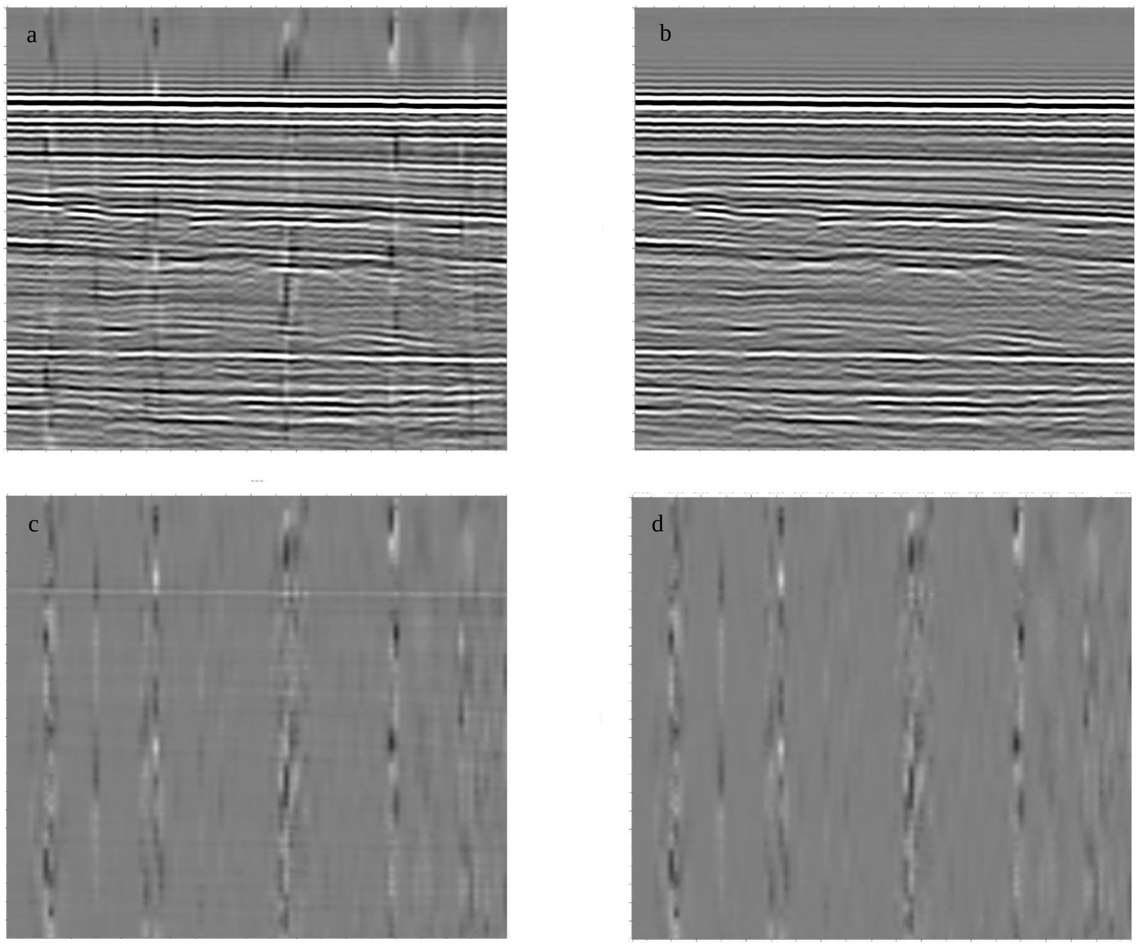

Figure 2: A common channel gather for a line with low swell noise: (a) input, (b) denoise output, (c) raw noise model, (d) noise model after BTN add-back. This demonstrates the effectiveness of our proposed workflow, which uses the supervised ML method to predict the noise model in one domain, and the self-supervised ML method to extract signal leakage from the noise model in another domain.

The coherent signal is then subtracted from the noise record to preserve the signal.

ML Denoise

The common-shot ML denoise flow is described in detail in Brusova *et al.* (2021). Before and after shooting, noise records contain recorded ambient noise from the cables. The ML denoise model is trained with input and target patches in a supervised manner. The input patches are built by combining random patches from the measured noise records and the processed, cleaned seismic records. The target patches (labels) are only patches from the measured noise records. The model is a CNN with a UNet-like architecture. The only difference between this abstract and the original paper is a significant increase in the training data (particularly noise records) to improve the robustness and diversity of the noise that the network can tackle. Locations of training data can be seen in Figure 1.

A Machine Learning Approach to Denoising and Signal Damage Reduction

Blind Trace Network

Laine *et al.* (2019) proposed an efficient blind-spot network in which they shift the receptive field of the convolutional network. To achieve this, all layers that modify the receptive field must be modified (e.g., convolutional, and upsampling/downsampling layers) but not those that do not modify the receptive field (such as 1x1 convolutions and concatenation layers).

To create a convolutional layer with a one-sided receptive field, we first pad the input on the side we want to restrict by half the kernel size, then perform a conventional convolutional layer, after which output has values on the opposite side cropped by half the kernel size. In the blind-spot network, a similar approach of padding and cropping is used to modify upsampling and downsampling layers to restrict the receptive field.

Rather than have four sub-models with operators acting with different receptive fields, Laine *et al.* (2019) add a rotation layer and rotate the input model to only include information to one-side of the output pixel; at the end a shift of 1 is performed and a series of 1x1 convolutions performed before estimating the output trace. A minor adaption can be made for the blind trace network, only using two rotations corresponding to receptive fields orthogonal to the time axis. Other than modifying the rotation and unrotation layers, the model is unchanged. Except for the above-mentioned changes, the model is a 5-layer UNet-like architecture.

Blind Trace Network for Signal Preservation

As stated above, this paper's novel contribution is applying the blind-trace network for signal preservation. The supervised ML Denoise flow is first applied in the common-shot domain. The noise model is then resorted in the common-channel domain. This is done to randomize the underlying shot-specific noise (in a trace, not pixel, sense) while the signal is expected to remain coherent. As this is a self-supervised technique, a project/line-specific model could be developed. To facilitate the application to production, the ML Denoise model was applied to a set of test lines, and these were used to train the blind-trace network. This single model was then applied for all lines, simplifying the production denoise flow. The model was implemented in pyTorch (Paszke *et al.*, 2019). Random cropping and random flip in the trace dimension were applied to improve the generalizability of the resulting model. The model's output contains the coherent signal energy not wanted as part of the noise model. The output can be added to the signal.

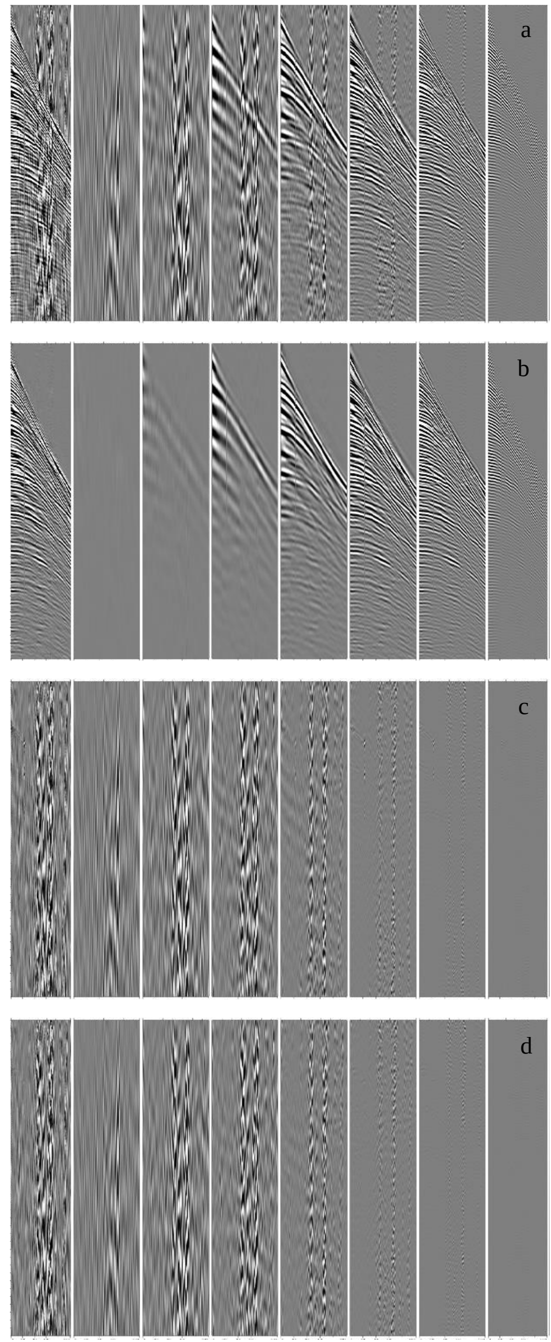

Figure 3: Octave panels for a shot record with high noise. From left to right panels are: Full bandwidth record, 0-0-2-4 Hz, 0-2-4-8 Hz, 2-4-8-16 Hz, 4-8-16-32 Hz, 8-16-32-64 Hz, 16-32-64-128 Hz, 32-64-128-256 Hz. (a) is input data, (b) data after denoise, (c) ML noise model (d) ML noise model with BTN signal removal.

A Machine Learning Approach to Denoising and Signal Damage Reduction

Results

The outlined technique was applied to a large regional project combining eight different legacy surveys from 1989 to 2019. Conventional denoise methodologies are challenging to apply in such a setting due to the variation in survey acqusition setup, conditions and resulting variability in noise levels requiring significant manual parameter optimization. The older surveys often have worse signal-to-noise at low frequencies due to shallow tow depths. Those surveys originally appled a low cut filter to discard frquencies up to 3-5Hz. Today, those frequencies are important to current techniques such as quality broadband imaging and FWI model building.

Different lines will be used to highlight different aspects of the proposed ML flow. The same flow was applied in all cases. Figure 2 shows a common channel gather windowed near the water bottom on a line with low noise. Figure 2(c) shows the result of the supervised ML denoise model. While the noise is identified, a weak imprint of the signal can be seen, particularly at the water bottom. This has no material impact on qualitative imaging but is of potential concern for quantitative processes such as AVO and, as such, is undesirable. Figure 2(d) shows the result of the add-back based on the Blind Trace Network.

Figures 3 and 4 are from a line that has a higher level of swell noise. Figure 3 shows octave panels for the first 100 channels and first 2 seconds as this is the window where signal leakage is most apparent. Octave panels are a key QC and highlight a number of interesting features. The primary takeaway is that the denoise flow results in clean records for all frequency scales. Looking at the noise records it is clear that there is some apparent signal leakage into the raw ML denoise record that is particularly apparent for near offsets (first 20 channels) at the waterbottom (around 0.3s). The primary damage is clearest in the 8-16-32-64 Hz and 16-32-64-128 Hz panels.

Another key observation from Figure 3 is the improvement in the 0-2-4-8 Hz panel. In the input the signal is well below the noise level. The final output shows a significant improvement. To further highlight this capability stacks have been generated in this bandwidth. In Figure 4, the stack of the raw data shows that the coherent energy beneath the waterbottom is overwhelmed by noise. The bottom figure shows the stack after the combined workflow. The stack of the denoised data reveals coherent events. Marked improvements are seen in the top half of the section (representing about 5-6 s) . The ability to effectively remove noise in the low frequencies while preserving coherent signal provides a strong foundation for other processing-steps, particularly deghosting and FWI.

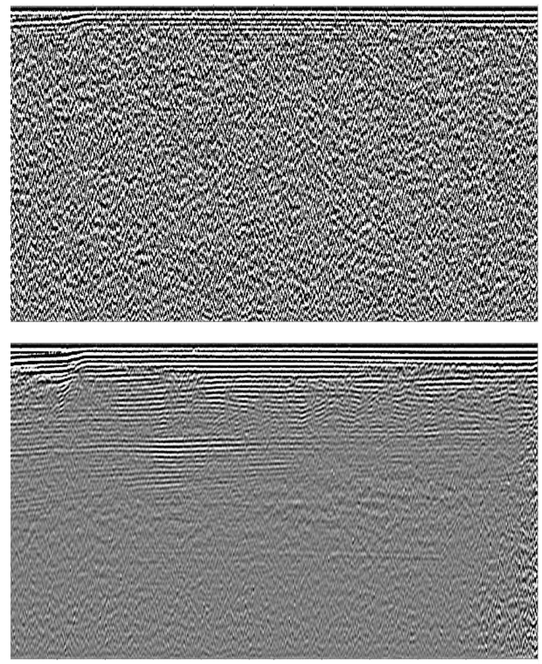

Figure 4: Stacks for 0-2-4-8Hz bandwidth data. Top stack is using raw input, bottom stack is using data after combined ML denoise flow.

Conclusions

A novel workflow combining supervised ML noise prediction and self-supervised ML signal addback has been shown to be effective as a leakage-free noise removal tool on multiple seismic lines. Results have demonstrated the efficacy of a generalized common-shot ML denoise model with the ability to recover valuable low frequencies on different vintages of streamer data. In areas of low noise, some signal leakage can be seen. We have demonstrated the ability of the self-supervised blind-trace network to identify the leaked signal. This combined approach applying two different models in two distinct domains leads to consistently high-quality results across data with a wide range of noise levels.

Acknowledgments

We thank TGS management for supporting publication of this abstract and for use of the data.