Applications of Deep Learning-based Deghosting in Marine Seismic Data Processing

Roald van Borselen[1], Ivan Vasconcelos[1], Sebastian Holland[1] , Rob Telling[1]

1. *Shearwater Geoservices*.

Summary

We demonstrate the potential of using Deep-Learning (DL) based de-ghosting methods in marine seismic data processing. We use a field-informed synthetic modelling strategy to create a diversified training data set, and account for acquisition uncertainties such as variable depth streamer profiles, wave action and variable surface reflectivity both in 2D and 3D applications. We present an application of source- and receiver deghosting on a UHR variable-depth streamer configuration, a 3D UHR example, and a field data application.

Keywords: Deep Learning, Deghosting, UHR, Wavefield decomposition

Applications of Deep Learning-based Deghosting in Marine Seismic Data Processing

Introduction

The presence of ghost reverberations – simply known as ghosts - deteriorates the marine seismic image, as ghosts constructively and destructively interfere with the primary wavefields. In Ultra-High Resolution (UHR) applications, receiver ghosts can be particularly problematic because wave-action and streamer-motion can cause a variable depth profile that will hamper the reliability of standard de-ghosting processing methods.

In geophysical literature, many have proposed solutions to address the problem of streamer depth variability and sea-surface time-variant roughness. Using the wave equation, one can derive operators in a transform domain that aim to generate ghost-free wavefields (e.g. Riyanti et *al*., 2008 and Grion et *al*., 2016). However, such methods require detailed information about receiver depths along the recording cables, the state of the water surface and the reflectivity of the free surface. This approach often leads to a compute- and user-intensive process to estimate the optimal parameters to obtain satisfactory results. An alternative to wave-equation-based methods is to use deep learning (DL), e.g. Baardman et al. (2022). Here, pairs of input data and desired outcome (labelled data) are used to train a network (the training stage) to apply an operator to unseen data in the inference stage. When relying on supervised learning, an important consideration in DL-based processing is how to train the network, using field data, synthetic data, or a combination of both.

In this abstract, we further explore the applicability of DL-based de-ghosting using synthetically trained networks. We discuss a strategy to ensure that synthetically created training data mirrors the field data as close as possible. We present a field data example, a synthetic UHR application of source- and receiver deghosting for a variable-depth streamer configuration and a 3D UHR example.

De-ghosting by supervised field-informed deep learning

In DL-based seismic processing, a chosen network is trained to learn a particular processing step, after which the trained network can be applied in inference mode to perform the said processing application to previously unseen data. It is well recognized that the quality of the DL-based processing results depends directly on what data are used to train the network. Using field-related information (e.g., well logs), input data can be chosen such that the training data captures the relevant characteristics the data that will be used in the inference stage – i.e., to potentially avoid the case of inference data being out of distribution relative to the training data priors. As such, there are several challenges to creating the required pairs of input-label data. One aspect is that the input field data are assumed to exhibit specific characteristics related to the processing application at hand, whereas other aspects are assumed to be constant or fully known. In the DL-application of deghosting, we aim to train a network that only “deghosts” the data, i.e. focuses on this one task. Using variably feathered input data would result in a network that tries to de-ghost and regularize the data simultaneously, which is generally undesirable. Secondly, to create the labelled data, a deterministic processing solution will have to be applied to some input data. Such a method will suffer from the fact that the input data never fulfils the perfect conditions for the deterministic solutions to work. Furthermore, there are intrinsic theoretical- and implementation assumptions in the deterministic method that will generally not be met by the field data. Another drawback of using field data for training purposes is that it may be time-consuming and compute-intensive to generate the labelled data.

An alternative to the use of field data is to train a network using synthetic data. Using synthetic modelling methods such as analytical methods, high-density finite-difference modelling or spectral element modelling, high-quality pairs of input-labelled data can be generated that provide high-fidelity examples of the desired output of the seismic processing step to be learned by the network. However, the obvious challenge with this approach is that the synthetic data may not resemble the inference data very well. To overcome this, we propose a field-informed modelling strategy that results in training data that mirrors the field data as closely as possible. This can be achieved in several ways. Firstly, when modelling the data, geological information, e.g. well-log-based models are to be used to describe the field data conditions as much as possible. Multiple geology models can be used to create robustness through generalization and data diversity. Secondly, modelling wavelets can be chosen to incorporate the effective source signatures deployed in the field, or directly extracted from acquisition data. Multiple wavelet variations can be used to increase a statistical significance Furthermore, one can rely on temporal & spatial sampling templates that match those of the inference field data. For the application of de-ghosting, we model a variety of receiver depth profiles to accommodate variable depth cables. Incorporating source- and receiver directivity effects into the modelling is also possible. During the training stage, we observe DL diagnostics such as training-and-validation RMS errors, and convergence rates to monitor model performance. This helps avoid overfitting, evaluate convergence and assessing computational efficiency in training.

In this abstract, the DL application of seismic deghosting is considered. To model the training data, we use an analytical method based on the Critical Reflection Theorem (CRT) (Fokkema et al., 1987) and we model the scattered pressure wavefield and the corresponding ghost-free data. In all examples shown, a 12-layer deep feed-forward Convolutional Neural Network (CNN) was utilized to train the data (Goodfellow et al., 2016).

Example 1: Synthetic Ultra-High Resolution (UHR) application

In this example, we consider DL-based de-ghosting in UHR marine seismic data processing. Applications such as the geotechnical assessment of offshore wind farm sites often require a detailed understanding of the properties of the near surface below the sea bottom. This is where 2D or 3D UHR seismic data are acquired with increased spatial and temporal rates over conventional seismic, e.g. 1 meter spacing and 0.1 millisecond, respectively. A high-frequency seismic source is typically towed at shallow depth, e.g. 0.25-0.75m and can emit energy up to 4000-5000 Hz. The wavefields are recorded with hydrophones positioned along streamers that are typically 50-150 m long and towed at shallow depths, e.g. 0.5-5m. Unlike exploration seismic, receiver positioning is often considerably more uncertain in UHR acquisition.

To train the network, data were generated for three different geological models, at different (constant) receiver depths of 0.3, 0.5, 0.7, 0.9, 1.1m. The cable length is set to 100 meters, with a hydrophone spacing of .5m and a sampling rate of .5ms. The peak wavelet frequency used in the modelling is 2500hz. For testing the method on the variable-depth inference data, we generate data using a V-shaped receiver depth streamer, starting at 0.4 meters at zero offset, linearly descending to 0.9 meters at 50-meter offset, and ascending linearly back to .4-meter depth at 100-meter offset. The source depth is set at .5 meters.

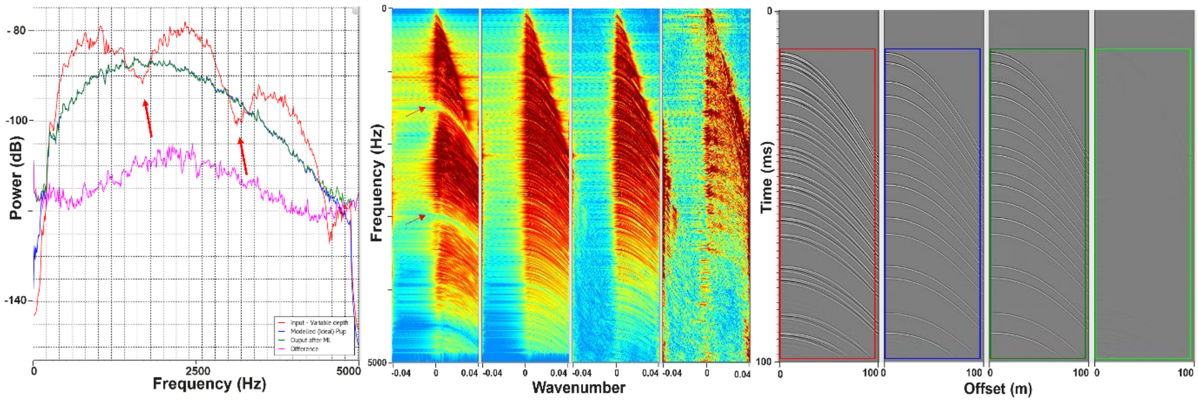

Figure 1: Spectra for the V-shaped 2D UHR data showing the input data (red), labelled data (blue), the DL-based source- and receiver deghosted data (dark green) and the difference (purple), the corresponding f-k spectra (centre) and the respective shot gathers (right).

Figure 1 shows the results for V-shaped variable depth cable data. On the left, the spectra are shown for the input scattered wavefield (red), the modelled (ideal) source- and receiver ghost-free upgoing wavefield (blue), the result after DL-based source- and receiver de-ghosting (dark green) and the difference (purple). In the centre, the corresponding *f-k* spectra, and on the right the respective gathers. The results show that the source notches at 1500Hz and 3000Hz (red arrows) and variable receiver notches are well reconstructed after DL-based de-ghosting. Note the very accurate correspondence between the modelled and predicted receiver-ghost free data, both displayed at reference depth .65m, where the observed difference is more than 20dB down in power.

Example 2: 3D DL-based Simultaneous Source- and Receiver deghosting

3D marine seismic exploration requires de-ghosting solutions that can account for the 3D characteristics of the measured wavefields. To accomplish this, there are two different strategies that currently prevail. Firstly, a full 3D deterministic wave-equation based solution can be used. This requires a fully sampled input data that fulfils the sampling criteria, both in in-line and cross-line. A well-known drawback is that the large streamer spacing in 3D acquisition is difficult to overcome and another is the high computational burden associated with such approaches. Another strategy is using a streamer-wise approach, where various approximations are employed to accommodate for 3D effects.

DL-based de-ghosting provides a potentially attractive alternative to the two strategies above. By relying on appropriate pairs of input and 3D labelled deghosted data, a network can be trained to produce a 3D deghosted result, for each streamer of acquired data. There are no 3D approximations, only the drawback of ensuring that training data remain statistically representative of the target data.

To demonstrate the effectiveness of such an approach, we modelled 12 streamers for 3D UHR data, again using the CRT method. Figure 2 shows the results for an outer streamer at 75m crossline offset. On the left, the spectra are shown for the input scattered wavefield (red), the modelled (ideal) source- and receiver ghost-free upgoing wavefield (blue), the results after DL-based source- and receiver deghosting (dark green) and the difference (light green) between the latter two. The corresponding *f-k* spectra are in the centre and on the right their respective gathers. The results show that the source notches (black arrows) and receiver notches (red arrows) are well reconstructed after DL-based deghosting. Note that the observed difference between the labelled data and the predicted data is more than 20 dB down in power.

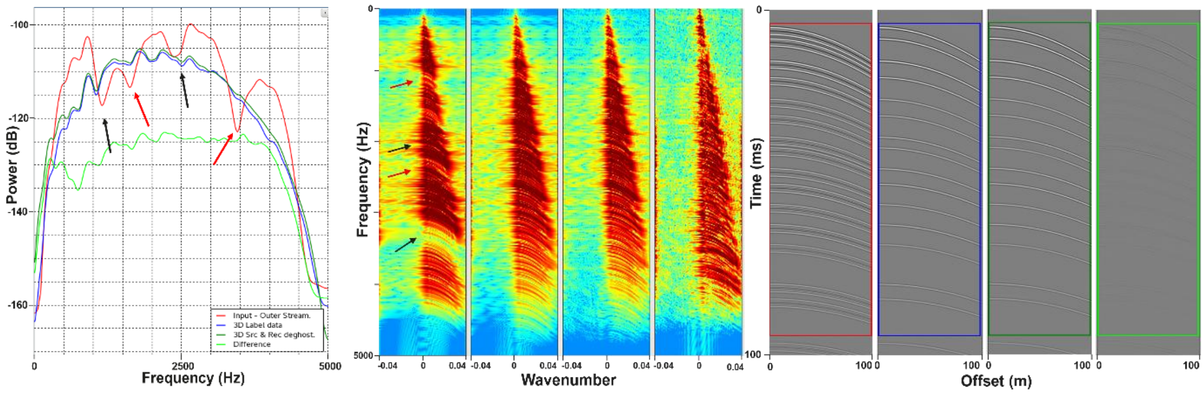

Figure 2: Spectra for an outer streamer of 3D UHR data showing the input data (red), labelled data (blue), the DL-based source- and receiver deghosted data (dark-green) and the difference (light-green), the corresponding f-k spectra (centre) and the respective shot gathers (right).

Example 3: Field data application

The final application is from the Link survey acquired in Namibia. Shot gathers from an outer streamer at 500m cross-line distance, consisting of 240 traces with 801 samples per trace at 12.5m hydrophone spacing were processed. To train the network, three interval velocity profiles with density estimates were created. Different source signatures were used that were derived from the data. To create diversity of the unknown receiver depths, data were modelled at five depths, around the nominal depth of 15 meters. The training data were modelled with the CRT method, where the exact field acquisition parameters were used. Figure 3 shows the spectra of an arbitrary gather before (red) and after DL-based receiver deghosting (blue) and after source- and receiver deghosting (dark-green). Note the excellent recovery of the respective notches at around 50Hz, caused by the nominal streamer depth of 15 meters, and at 93Hz, caused by the source towed at 8 meters The *f-k* spectra of the respective results are shown in the centre. Here the recovery of the receiver notch (black arrow) and the source notch (red arrow) is again clearly visible. On the right-hand side, the seismic gathers are displayed. It is remarked that the training of the network took around 30 minutes on a single GPU, whereas the field data were processed at around 8 seconds per shot gather, with 240 traces and 801 samples per trace.

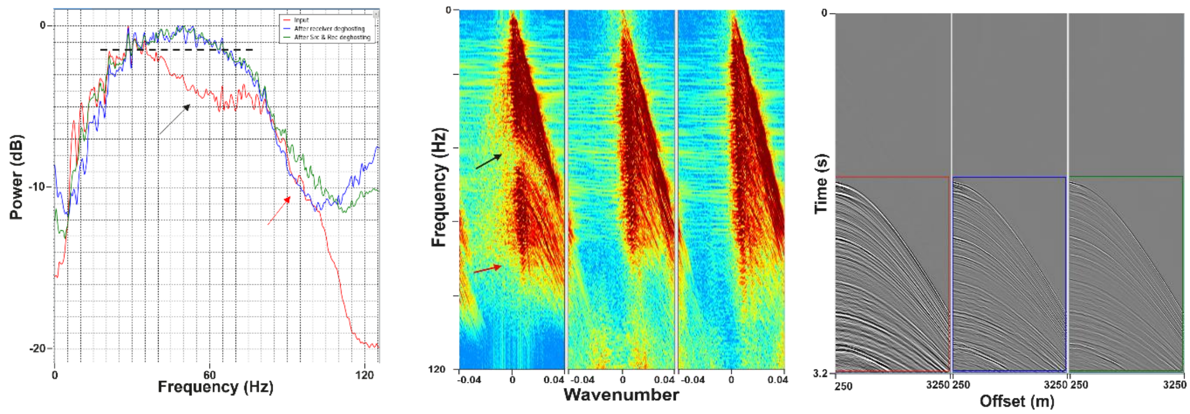

Figure 3: The spectra of the shot gather before deghosting (red), after DL-based receiver deghosting (blue) and DL-based source- and receiver deghosting (green), the corresponding f-k spectra of the gathers data (center), and the respective shot-gathers (right).

Conclusions

We demonstrate the potential of Deep-learning based de-ghosting methods in marine seismic data processing. Using a field-informed strategy to create a diversified training data set that resembles the inference field data applications as much as possible, acquisition uncertainties such as variable depth streamers, wave action and variable surface reflectivity can be accounted for in 2D and 3D applications. Satisfactory results are obtained efficiently with minimal user-interaction, outside of that required to engineer the appropriate training data. This is particularly useful in fast-track processing, aimed at delivering a data volume for early interpretational work and/or velocity model building.

Acknowledgements

We thank Searcher Seismic for permission to show the field data example.

References

Baardman, R.H., Yoo, J. & Hegge, R.F., (2022). Supervised Learning Approach for Multi-Component Deghosting with a Focus on Shallow Water Data. First Break, Volume 40, Issue 12, p. 51 – 57.
Fokkema, J.T. & Ziolkowski, A.M., (1987), The critical reflection theorem, Geophysics 52: 965-972.
Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
Grion, S., Telling, R. Holland, S., 2016, Rough sea estimation for phase-shift deghosting, SEG Technical Program Expanded Abstracts, p 5129-5133.
Riyanti, C. D., Borselen, R., Berg, P. M., & Fokkema, J. (2008). Pressure wave-field deghosting for non-horizontal streamers, SEG Technical Program Expanded Abstracts: 2652-2656.