Introduction

The higher signal-to-noise (S/N) at low frequencies provided by seabed seismic acquisition presents an opportunity to learn, in a supervised machine learning fashion, a function for artificial bandwidth extension that can be used to enrich geophysical datasets missing low frequencies. The missing ultra-low frequency band (below 3 Hz) for acquisition systems such as towed streamer are known to considerably limit the accuracy of subsurface products obtained from full-wavefield inversion (FWI). Efforts made to-date by the industry to address this lack of low frequencies have focused on (1) stronger signal from more advanced source types and deeper tow (broadband), (2) more sophisticated broadband preprocessing techniques and regularization schemes to solve the inverse problem, (3) recently, investigating the possibility of manufacturing low frequencies from examples using deep learning.

Supervised machine learning to construct low frequencies aims to train a neural network based on many examples containing both the high frequencies (input) and the low frequencies (target). A fundamental assumption is that, once the network is trained, it will be applied on data (test set) with features that are a subset of those contained in the training data. Many methods proposed lately for low frequency extrapolation using deep learning have shown promising results (Hu et al. 2019, Jin et al. 2018, Ovcharenko et al. 2019, Sun et al. 2019). A challenge however is the learning based on synthetically-generated labels (low frequencies) which must limit generalization when field data exhibit complex features. In this regard, learning from realistic seabed seismic data presents the advantage of higher-quality labels and therefore potential for higher-fidelity bandwidth extension.

Low frequency extrapolation can naturally be attempted in the frequency-wavenumber (f-k) domain. By considering f-k slices to be independent, approaches relying on purely feedforward neural networks miss taking advantage of correlations between adjacent frequencies (Figure 1). One way to account for these correlations is by looking at a set of frequencies in a given frequency band as a sequence, i.e. a collection of objects in which correlations exist. From that insight, we pose low-frequency extrapolation as sequence prediction problem, with the goal of mapping an input sequence of high frequencies (e.g. 3-8 Hz) into an output sequence of low frequencies (0.5-3 Hz).

Sequence Prediction

Seq2Seq, short for Sequence to Sequence, is a generic neural-network-based framework for learning relationships between input and output sequences allowed to vary in length and nature. Among several other important applications, Seq2Seq has shown remarkable performance in natural language processing and neural machine translation (Cho et al. 2014, Sutskever et al. 2014). Many geophysical datasets reveal themselves in sequential form: for example, the f-k representation of seismic data can be seen as a sequence of f-k slices (Figure 1), each slice sharing correlations with its neighbors.

![Figure 1 FK representation of seismic data as a sequence of frequency slices](image)

One way to implement a Seq2Seq model is via an encoder-decoder architecture. The task of the encoder is to operate over an input sequence of high frequencies, one element at a time, and encode correlations existing between its elements into a high-level vector representation. The task of the decoder is to predict a sequence of low frequencies conditioned on the input sequence of high frequencies and the high-level vector representation obtained from the encoder. The key in both the encoding and decoding processes
is achieving a “fine-grained” learning of correlations between the sequence elements. The encoder and decoder can be trained jointly on many pairs of high and low frequencies extracted from ocean-bottom nodes or cables. By exposing the model to many examples, the high-level vector representation estimated by the encoder can be thought of as summarizing the complexity of the medium (geology) the network has been exposed to.

There is a chance to have the trained network generalize well as long as it is applied on data “similar enough” to the training input. Such scenario can happen in the case, for instance, of towed-streamer data acquired within similar shot coverage as the ocean-bottom nodes. Owing to source-receiver reciprocity, a very sparse set of ocean-bottom nodes deployed randomly may prove sufficient for training. Furthermore, a carefully crafted preprocessing pipeline composed of broadband processing (Aharchaou and Neumann 2019), wavefield redatuming, multiple attenuation, and other tools can be employed to reduce differences between seabed seismic and towed-streamer data.

We can choose among many possible neural networks to form the encoder and the decoder. Recurrent neural networks, e.g. long short-term memory (LSTM), are a natural fit for dealing with sequences. Stacking multiple LSTMs can give a higher representational power at the encoding/decoding levels. Significant performance uplift can further be achieved through “attention mechanisms” allowing the network to learn where to attend in the input f-k sequence for each low frequency-wavenumber sample in the output sequence.

Field data example

We test the feasibility of learning artificial bandwidth extension from seabed seismic on field data consisting of 1464 ocean-bottom nodes. Each node records traces from ~300K shots, with wide shot coverage and good S/N down to 1 Hz. We randomly select 10% nodes (Figure 2) to form the training data, made of pairs of high frequencies (3-6 Hz) and low frequencies (1-3 Hz), amounting to ~45M training samples. Training on only 10% nodes, we ask: “How accurately can the trained Seq2Seq network extrapolate low frequencies given high frequencies from the 90% nodes not part of training?” A benefit from such test is the availability of “ground truth” low frequencies to compare against. More importantly, we are interested in how FWI responds to the manufactured low frequencies.

Figure 2 Ocean-bottom node map (1464 nodes): red nodes (10%) used for training, black nodes (90%) used for testing. Each ocean-bottom node is exposed to wide shot coverage.

Figure 3 shows an example of manufactured low frequencies on a node gather from the test set. Even with an advanced noise suppression workflow, the “ground truth” ultra-low frequencies still seem dominated by noise for deeper events, highlighting one challenge about learning from field data. The manufactured low frequencies seem considerably less noisy, exhibiting an excellent phase/amplitude
match with the original low frequencies for many strong-amplitude events, but also lack signal in parts of the gather challenged by weak S/N. Depending on the objective function used by FWI, the phase information may turn out to be more important than the amplitude information.

**Figure 3** Manufactured low frequencies for a test node gather. From left to right: a) input high frequencies (3-6Hz), b) original low frequencies (1-3Hz), c) manufactured low frequencies (1-3Hz).

Figure 4 compares the response of FWI to the manufactured and original ultra-low frequencies. To enable more accurate updates in the salt mini-basin, a 2 Hz high-cut filter was applied on the input data. The difference between the two Vp models at convergence (iteration 19) indicates that there is an overall good match between the two solutions, except for updates in the salt flanks related to the challenge of accurately extrapolating the weak diffractions.

**Figure 4** FWI response to original and manufactured low frequencies. From left to right: a) starting Vp model, b) FWI Vp inverted model (iteration 19) based on original low frequencies (with 2 Hz high-cut filter), c) FWI Vp inverted model (iteration 19) based on manufactured low frequencies (with 2 Hz high-cut filter), d) difference between the two inverted Vp models.
Conclusions

The good S/N at low frequencies provided by seabed seismic allows the learning of an artificial bandwidth extension function that can be used to enhance geophysical datasets missing low frequencies. By training on a very sparse OBN set, which would be considerably cheaper to deploy than the full set, the trained Seq2Seq model can be applied in a “narrow generalization” sense to enhance towed-streamer data over the same region. Never a substitute to real low frequencies, the manufactured low frequencies may nevertheless prove useful in the inversion of subsurface properties, the same way interpolated traces are useful in reducing migration swing and enhancing seismic image quality. The frequency-wavenumber representation of seismic data allows taking advantage of correlations between neighbouring high-frequency f-k slices to provide more “context” to extrapolate low frequencies. Although we focused on the low-frequency band, the same framework can be used to extend the high-frequency band by learning, for instance, from shallow-tow streamer acquisition. Such mapping may be useful in manufacturing high frequencies for “broadband” deep-tow datasets.

Acknowledgements

We thank ExxonMobil for allowing us to publish this work. We also thank CGG for permission to use the data.

References


