Training Strategies for Supervised Deblending - DSA Data Example

Rolf Baardman[1] and Rob Hegge[1]

[1]Aramco Overseas Company BV.

Abstract

Objectives/Scope:
In this abstract we discuss practical considerations regarding a supervised machine learning (ML) based approach to deblend seismic data from a Distributed Source Array (DSA) acquisition. The blended DSA survey consists of 24 sources (12 high- and 12 low-frequency sources) that were deployed independently to significantly reduce turn-around time. The success of the supervised ML-based deblending method depends primarily on the quality of training data. Two important factors to consider are user-friendliness (the effort it takes to obtain training data) and the level of similarity between training and unseen data to ensure generalization. We discuss various approaches to obtain training data and illustrated these on the DSA data. To select the best strategy we consider factors like turn-around time, quality of conventional deblending methods, data availability, and user-friendliness.

Methods, Procedures, Process:
One strategy to obtain training data is to model and computer-blend data gathers. The main challenge of this strategy is to ensure the generalization; it is a cumbersome and case-specific process to obtain modelled data that matches the characteristics of the unseen data. This is a strong requirement to ensure good deblending results. Another training strategy is to use field data with characteristics similar to the unseen data. A first option for this strategy is to apply an existing deblending algorithm to a small subset of the unseen data. Once trained, the rest of the data can be processed very quickly. A disadvantage is its reliance on existing deblending methods; running it could still be user-intensive and the quality could bias the training process. A second option is to use unblended field data and artificially blend it. The main advantage is its independence of conventional deblending methods that could bias the deblending performance. The viability of this option relies on data availability; it requires an unblended survey that was shot in the vicinity or, if logistics allow for it, to shoot additional unblended lines specifically for training purposes. Both options for training a neural network using field data are illustrated on the DSA data.

Results, Observations, Conclusions:
The ML-approach using the first field data option performs as good as the deblending method used to obtain the training data itself. The second option uses unblended data from a survey in the vicinity of the DSA survey. Results were very promising and in particular the signal preservation was excellent. Performance was slightly better compared to the result of the first option where some leakage was visible due to the bias created in the training. Overall, we can conclude that both ML-based deblending options perform very well and are at least on par with conventional methods. If the proper training data are or can be made available, the second option would be the preferred approach.