Submission code: 9P4VADY9P4

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Machine learning recovery of low frequency content for seismic data

Pablo Machado Barros (Cenpes-Petrobras), André Bulcão (Petrobras), Bruno Pereira Dias (Petrobras), Claiton Pimentel de Oliveira (Petrobras), Djalma Manoel Soares Filho (Petrobras), Gustavo Catão Alves (Petrobras), Tiago Ilipronti Girardi (Cenpes-Petrobras), Ubiratan Jose Furtado (Petrobras), Luis Fernando Mendes Cury (Cenpes-Petrobras)

SBGf Conference 18-20 NOV | Rio'25

Machine learning recovery of low frequency content for seismic data

Please, do not insert author names in your submission PDF file

Introduction

Full-waveform inversion (FWI) is the current industry-standard for determining velocity models from seismic data. This inversion process requires data with reliable low frequency content to avoid problems such as cycle-skipping. During the seismic processing workflow, we aim to transform our acquired seismic data as if it was acquired through a well-behaved known seismic source. In this work we aim to transform our seismic data from a given seismic source signature to one with a wider bandwidth on the lower frequency range through machine learning (ML). We then compare the FWI-obtained velocity models with both the original and transformed data.

Method and/or Theory

We use a *u-net* like deep neural network architecture that takes as input the seismic data acquired with a given seismic source signature and outputs the same data with a target source signature. The *u*-net architecture is characterized by a sequence of downsampling blocks, a bottleneck block, and a sequence of upsampling block. Each downsampling block has some convolution layers followed by a pooling layer, resulting in an output with more channels but in which each channel has a lower dimension than in the previous block. Each upsampling block has some convolution layer followed by a transpose convolution, resulting in an output with less channels but in which each channel has a higher dimension than in the previous block. The bottleneck block is simply a sequence of convolution layers. Equivalent downsampling and upsampling blocks are connected by skip connections. We choose a Ricker wavelet with peak frequency of 26Hz that reaches -20dB signal intensity around 6Hz and 53Hz as the seismic source for the input data. For the target data we choose as the seismic source a SEP wavelet with peak frequency of 26Hz that reaches -20dB signal intensity at around 0.6Hz and 53Hz. This way we tell the neural network to restrict its modifications to the range below 6Hz. We use synthetic seismic data from a streamer acquisition modeled on public and proprietary velocity models both as acoustic and elastic waves. Streamer data usually present low SNR in the low frequency range, leading this part of the data to be discarded. It is our aim that this range of data could be recovered through ML transformation creating a better basis for the seismic inversion.

Results and Conclusions

Our preliminary results show that a large amount of data is needed to avoid overfitting while allowing the ML model to give good results on unseen data. The research is ongoing, and we are currently training new models on a larger dataset. We intend to compare the resulting seismograms with the original data, and then show its impact further down the processing pipeline by comparing the velocity models obtained by the inversion of the resulting and original seismic data.