Source wavelet estimation with machine learning and generalized seismic wavelets

Ramon Oliveira Junior[1], Ramon C. F. Araújo[1], Samuel Xavier-de-Souza[2], Tiago Barros[2]. Gilberto Corso[3],

[1] Departamento de Física Teórica e Experimental, Universidade Federal do Rio Grande do Norte, Natal, RN, 59078-970, Brazil.
[2] Departamento de Engenharia de Computação e Automação, Universidade Federal do Rio Grande do Norte, Natal, RN, 59077-080, Brazil.
[3] Departamento de Biofísica e Farmacologia, Universidade Federal do Rio Grande do Norte, Natal, RN, 59078-970, Brazil.

Abstract

Source wavelet estimation from seismic data is an important task in the subsurface imaging effort. In this study, we intend to test a Machine Learning (ML) approach to obtain the wavelet parameters from seismograms. For the sake of simplicity, it is usually assumed that the Ricker wavelet is a good approximation for the source waveform. However, the attenuation caused by the propagation in heterogeneous media distorts the frequency content of the wavelet, which is an effect not represented suitably by the Ricker. To correct this phenomenon, we work with the Generalized Seismic Wavelet (GSW), which is a recently proposed generalization of the Ricker wavelet. The Ricker has just one parameter, the fundamental frequency. On the other hand, the GSW has two parameters: the same fundamental frequency and the order of the fractional time derivative of the Gaussian function; the latter controls the wavelet's shape and asymmetry in time domain. For a specific value of this additional parameter, the GSW becomes the Ricker wavelet. Using the seismogram as input, machine learning models are trained to output the generalized wavelet parameters that best explain the data. The source wavelet is then estimated as the GSW with the parameters provided by the model. The proposed methodology was tested and validated with synthetic seismic data, modeled using a 1D velocity model obtained from the Gato do Mato, an oil field located in Santos basin, Brazil. We propagate an acoustic wave using the following acquisition geometry: one source placed at the surface and 50 receivers placed at 2100 m in depth, spaced 100 m apart from one another in the horizontal direction. We performed one thousand numerical propagations using the GSW with different parameterizations as the source waveform. Furthermore, we randomly selected 70% of the data for training, and the other 30% for testing. Three ML algorithms have been considered for regression: Extreme Gradient Boosting (XGBoost), Bayesian Ridge regression and the Convolutional Neural Network. The preliminary results were satisfactory, achieving good accuracy in estimating the GSW parameters.