Improving the Retrieval of Offshore-Onshore Correlation Functions with Machine Learning

*Loic Viens¹, Tomotaka Iwata¹

1. DPRI, Kyoto University

The retrieval of reliable offshore-onshore correlation functions is critical to improve our ability to image subduction zones and to predict long-period ground motions from potential megathrust earthquakes. However, localized and persistent ambient seismic field sources between offshore and onshore stations can bias both the amplitude and phase information of correlation functions and generate non-physical spurious arrivals. In this study, we present a two-step method based on unsupervised learning to improve the quality of correlation functions calculated with the deconvolution method (e.g., deconvolution functions, DFs). For a given DF set calculated between two stations over a long time period, we first reduce its dimension using the Principal Component Analysis (PCA) and then cluster the features of the low-dimensional space using a Gaussian mixture model. We finally stack the DFs belonging to each cluster together and select the cluster that maximizes the symmetry between the anti-causal and causal parts and minimizes spurious arrivals. We focus on DFs calculated every 30 minutes over one year between offshore DONET stations located on top of the Nankai Trough, Japan, and 77 surrounding onshore Hi-net receivers. Our method allows us to remove spurious arrivals, which dominate the raw stack of the DFs over the year, and therefore to significantly improve the signal-to-noise ratio of the DFs. To demonstrate that both the amplitude and phase information of DFs is preserved with our method, we use them to simulate the ground motions from a $M_{\rm w}$ 5.8 earthquake, which occurred on April 1, 2016 in the vicinity of the KME18 DONET station. Results show that the long-period (4-10 s) ground motions from the earthquake can be accurately simulated. Our method can easily be used as an additional processing step when calculating DFs between offshore and onshore stations to improve the retrieval of reliable phase and amplitude information, and to remove potential spurious arrivals.

Keywords: Machine Learning, Ambient seismic field, Correlation function, Green's function, Long-period ground motions