Automatic signal classification for continuous seismic waveform records based on an unsupervised learning algorithm: Application to artificial noises and low-frequency tremors

*Yuki Kodera[1], Shin'ichi Sakai[2], Takahito Nishimiya[1]

1. Meteorological Research Institute, Japan Meteorological Agency, 2. Earthquake Research Institute, the University of Tokyo

Continuous seismic records catch various signals such as earthquakes, human activities, and instrumental noises. Extracting and analyzing these signals are important for earthquake early warning (EEW) operations. Using earthquake signals, we can evaluate the detection capability of the seismometer and calibrate the site amplification factor. Artificial noise signals tell us about the current data quality of the seismometer and the possibility of false alarm issuance. In addition to the EEW purpose, classifying natural and non-natural signals may contribute to the better understanding of geophysical phenomena around the observation network. In this study, we have developed an automatic classification algorithm for continuous records based on a machine learning method.

An unsupervised learning algorithm is employed in our approach because observed signals in continuous records vary depending on every seismometer. We assume that we do not have the knowledge of the recorded signals in advance and that labeled training data are not available that can be used for a supervised learning. The method is based on clustering in the frequency and time domains. First, features are obtained by computing running spectra with a 4-s time window and 0.1-s interval, applying a triangular filter bank with 10 bands. After that, the features are clustered in the 10-dimensional feature space using the k-means algorithm. Then, the spectral clustering algorithm is performed with the assumption of the transition matrix of a Markov chain. Ten time-series models are finally obtained and used for the classification.

We applied the method to continuous waveforms recorded at (1) Station E.JDJM, which is a seismometer of the Metropolitan Seismic Observation network (MeSO-net) located close to the subway and (2) a temporary ocean-bottom seismometer (OBS) to observe aftershocks of the 2004 M7.4 off the Kii Peninsula Earthquake (Yamazaki et al., 2008). For the record of Station E.JDJM, the algorithm discriminated earthquake signals and train noises from background noises. For the OBS record, low-frequency tremors with large amplitudes were automatically detected in addition to regular earthquakes. The method successfully classified continuous records of the MeSO-net station and OBS although those seismometers had unique observation environments and diverse recorded signals. This indicates the potential for the wide applicability of the unsupervised learning approach and the contribution of monitoring seismometer conditions and geophysical phenomena.

Keywords: Machine learning, Earthquake early warning, Low-frequency tremor, Automatic classification, MeSO-net, Ocean-bottom seismometer

