10th ACES International Workshop, Awaji Island, September 25-28, 2018

Application of Extreme Value Theory to a Matched-Filter Analysis

Hirano, S.[1], Toyomoto, Y.[1], Kawakata, H. [1], and Doi, I.[2]

1. *College of Science and Engineering, Ritsumeikan University, Shiga, Japan*
2. *Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan*

Relatively small foreshocks, aftershocks, swarms, and LFEs have been detected by using Matched-Filter (MF) analyses, which enables us to extract seismic signals from continuous noisy waveforms. In a traditional MF technique, a normalized Cross-correlation Coefficient (CC) between the continuous and template waveforms was calculated, and parts of the continuous waveform where CC with the template exceeds a threshold value have been regarded as the seismic signals. In this procedure, however, choice of the threshold is subjective, and misdetection could occur with an inappropriate threshold value if we do not understand the statistical properties of CC.

In this study, we discuss statistics of CC and compare theory to data for foreshocks prior to an M5.4 Nagano earthquake in Japan on June 30, 2011. Theoretically, we can show that the histogram of CC is expected to follow the normal distribution of zero mean and $1/N_{\mathrm{smp}}$ variance, where N_{smp} is the number of points of the template (e.g., $N_{\mathrm{smp}} = 500$ for 5 seconds data of 100 Hz) for arbitrary template waveforms if the continuous waveform is independent and identically distributed. However, we have found that data do not fit this expectation but are more fat-tail, which means that setting a threshold for CC likely to cause a misdetection. On the other hand, the maximum value of CC in every 5 seconds follows the Gumbel or Generalized Extreme Value (GEV) distribution that explains the rare emergence of extreme values. Hence, we can define an objective and quantitative indicator to detect seismic events by comparing the theoretical distribution and data histogram. However, we also find that, in some cases, fitting by the distributions with parameters based on *L*-moments does not explain the tail. Hence, we execute maximum likelihood estimation of the fitting parameters and discuss how we should regard outliers as seismic events.