



A k-mean characteristic function to improve STA/LTA detection

Jubran Akram^{*1}, Daniel Peter¹, and David Eaton²

¹King Abdullah University of Science and Technology (KAUST), Saudi Arabia

²University of Calgary

Summary

STA/LTA is commonly used to detect microseismic events because of its simple structure. Although computationally efficient, this algorithm becomes inconsistent when input waveforms have low S/N. To optimize the detection quality, we propose a modification to the conventional STA/LTA workflow. Instead of using amplitude or energy of the input waveform, we use a k-mean characteristic function which is more sensitive to signal variations relative to background noise. In addition, we also use the first and second order differences of input waveforms for k-mean characteristic function. Using a synthetic data that comprise of 11 microseismic events, we compare the proposed workflow with the conventional STA/LTA. For a constant background noise level, our results indicate that conventional STA/LTA is less responsive to and therefore, fails to detect two weakest events ($M_w = -3, -2.8$) despite a very low threshold value (2). However, STA/LTA as applied in our proposed workflow successfully detected each of 11 events at the cost of one false positive.

Introduction

Detecting microseismic events from the continuous data recordings is an important processing step. Although manual scanning can be done, it becomes highly inefficient when data spans several hours or days of continuous recording. In this case, automatic algorithms are used to detect microseismic events effectively. Two most common automatic event detection algorithms are short and long-term average ratio (STA/LTA) and correlation detector. Here we focus on STA/LTA algorithm (Allen 1978; Earle and Shearer, 1994; Withers et al. 1998), which has been used extensively in the earthquake seismology and in the petroleum industry because of its simplicity. In STA/LTA, average of a characteristic function such as absolute amplitudes or energy of the input signal is computed in two moving windows. The smaller window provides information on signal fluctuations whereas larger window is representative of background noise level. Thus, the ratio of these windows (STA/LTA) when exceeded by a user-specified threshold triggers a microseismic event detection.

Often, quality STA/LTA detection requires careful selection of window sizes and a threshold value. Ignoring these parameters, especially for low S/N waveforms, can result in either missing many events or recording too many (false positives). We aim to improve STA/LTA detection such that it detects weaker events with less false positives. For this purpose, we introduce a k-mean curve as a characteristic function, which is obtained by taking the mean value of k-nearest-neighbors (Schelter et al., 2006) of i^{th} sample in a moving window fashion. The STA/LTA that is obtained from the proposed characteristic function is less sensitive to the choice of threshold value.

Event detection

Let $S \in \mathbb{R}^{M \times 3}$ be 3C continuous recording at the i^{th} receiver level. In practice, waveform data are reduced to smaller windows ($\mathbb{R}^N \times 3$) with some overlap for visualization without computer memory limitations. Event detection on this data is then performed as follows:

Step 1: For each component, compute first and second order differences. Now we have three sets of 3C data: 1) Original 3C waveforms 2) first order difference of input 3C waveforms 3) second order difference of input 3C waveforms. Steps 2-6 are repeated for each set of 3C data.

Step 2: For each component, compute the Euclidean distance of each sample from the remaining samples

$$d_{ij}^{(p)} = |x_i^{(p)} - x_j^{(p)}|, \quad p = 1,2,3 \text{ and } i, j = 1,2,3 \dots, N. \quad (1)$$

Step 3: For each sample, find the k-nearest neighbors and compute their mean value to obtain a characteristic function (C)

$$C_i^{(p)} = \frac{1}{k} \sum_{q=1}^k d_{ij}^{(p)}(q), \quad (2)$$

where $d_{ij}(q)$ is the q^{th} nearest neighbor for i^{th} sample from j distance values.

Step 4: For a smooth characteristic function, apply a moving average filter of length k on C

Step 5: To combine characteristic functions from three components effectively, apply min-max normalization and stack

$$C_{i,norm}^{(p)} = \frac{C_i^{(p)} - \min(C_i^{(p)})}{\max(C_i^{(p)}) - \min(C_i^{(p)})}. \quad (3)$$

$$C_i = \frac{1}{3} \sum_{p=1}^3 C_{i,norm}^{(p)}. \quad (4)$$

Step 6: Compute STA/LTA on the characteristic function from Eq. 4.

$$\left(\frac{STA}{LTA}\right)_i = \frac{\frac{1}{NS} \sum_{j=i-NS}^i C_j}{\frac{1}{NL} \sum_{j=i-NS-NL-1}^{i-NS-1} C_j}, \quad (5)$$

where NS and NL represent the length of short and long-term windows, respectively.

Step 7: Stack STA/LTAs that are obtained from 3C waveforms and their differences.

Step 8: Apply a threshold criterion on the stacked STA/LTA to detect microseismic events.

Application on synthetic data

To evaluate the proposed workflow, we generated synthetic waveforms by using far-field displacement equations for numerous microseismic events of descending magnitudes. The minimum and maximum magnitudes used in this synthetic modeling were -3 and -1, respectively. A total of 11 events were generated with interval of 0.2. All microseismic events shared the same source mechanism (moment tensor). To mimic a real experiment, we also included same level of white gaussian noise. Figure 1 shows the 3C waveforms for a microseismic event ($M_w = -3$) with and without the addition of noise. For this event, the background noise level consumes the signal on x- and y-components. However, signal is slightly visible on the z-component. The distribution of all events in the background noise series is chosen through random sampling. Figure 2 shows entire 3C waveforms with Gaussian noise that were used to evaluate and compare the proposed algorithm with conventional STA/LTA. Stronger events are clearly visible as compared to smaller magnitude events. However, a zoomed view between 4s and 5.5s on the y-component (rectangle) also shows two microseismic events of magnitudes -2.0 and -1.6.

Figure 3 shows the STA/LTA detection for three different scenarios (S1-3). STA/LTA were generated using same window sizes in all scenarios. In S1, STA/LTA curve is generated from the stack of amplitude envelopes of 3C input waveforms. A very low threshold value (2) detected 9 out of 11 events successfully without any false positives. Two missed events were the smallest magnitude events (-3 and

-2.8). In S1, lowering the threshold any further can produce many false positives as the used threshold is already near background level. The performance of STA/LTA is similar in S2 in which we stacked the STA/LTA from each of the 3C waveforms, and their first and second order differences. However, using the proposed workflow (S3), we successfully detected each of 11 events along with a false positive by using a threshold value of 10. This false positive may be removed when STA/LTA detections are correlated across the receiver array.

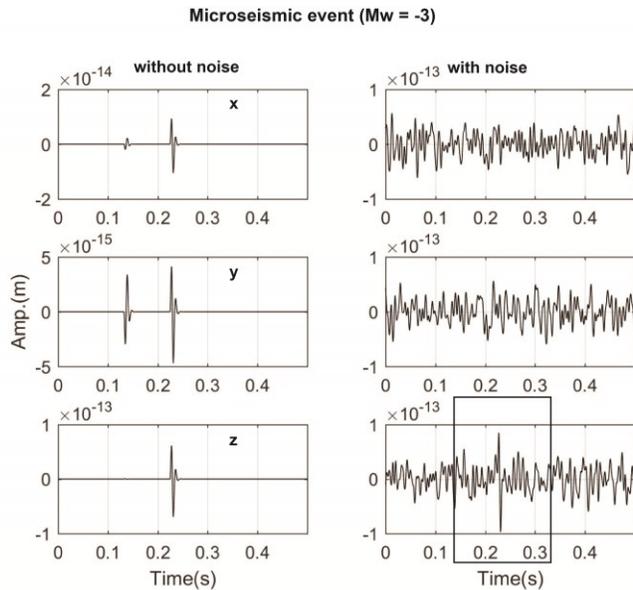


Figure 1: An example of a synthetic microseismic event of Mw = -3. Being the weakest event, the addition of Gaussian white noise consumes the signal in both x- and y-components. However, S-signal remains slightly visible in the z-component as shown by a rectangle.

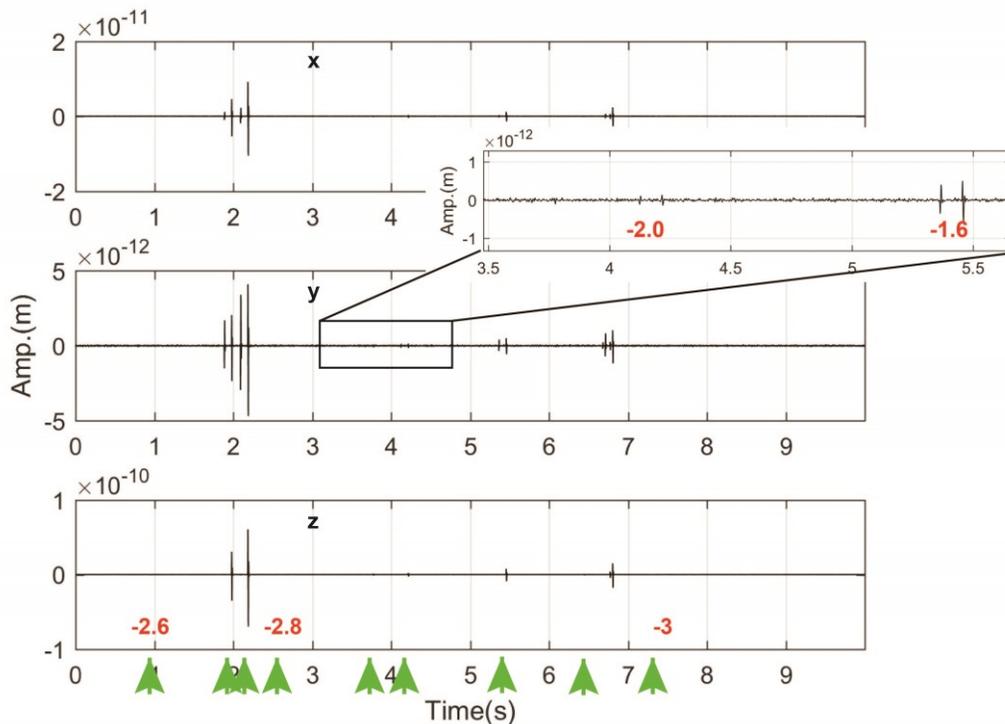


Figure 2: 3C noisy waveforms used to evaluate and compare the proposed algorithm with conventional STA/LTA. White Gaussian noise level is constant. Several microseismic events of magnitudes that range between -3 and -1 were placed randomly (green arrows) in the background noise series. A zoomed view on the y-component (rectangle) also shows weaker microseismic events.

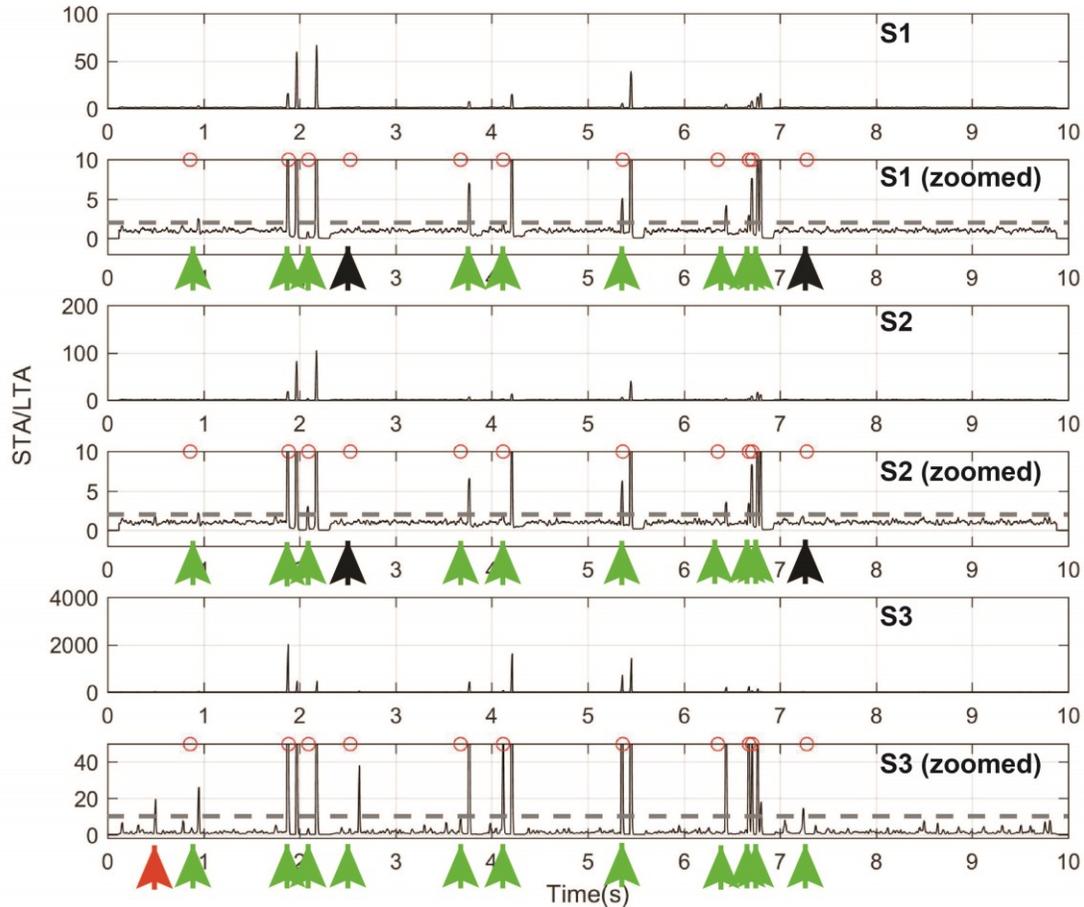


Figure 3: STA/LTA detection for three different scenarios (S1-3). S1 shows the STA/LTA detection when stack of amplitude envelopes of 3C waveforms are used as input. In S2, we stacked the STA/LTA from each of the 3C waveforms, and their first and second order differences. However, in S3, we further compute the k-mean curve of the 3C waveforms and their differences before using them as input for STA/LTA algorithm. The STA/LTA curve in S3 is the stack of these 9 STA/LTA curves. Dashed horizontal line shows the detection threshold (S1, S2 = 2, S3 = 10). Green arrows indicate successful detection, and red arrows show false positive whereas black arrows indicate missed detection.

Conclusions

We proposed a modification to the conventional STA/LTA workflow. Using first and second order differences of the 3C input waveforms in the workflow increases the sensitivity of STA/LTA to the amplitude fluctuations. Also, including k-mean characteristic function of the input waveforms and its differences further enhances the sensitivity of the STA/LTA to the amplitudes from weaker events. More weaker events can be detected at the cost of a lower number of false positives.

Acknowledgements

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST).

References

Allen, R., 1978, Automatic earthquake recognition and timing from single traces: *Bulletin of the Seismological Society of America*, 68, 1521–1532.

Earle, P. S., and P. M. Shearer, 1994, Characterization of global seismograms using an automatic-picking algorithm: *Bulletin of the Seismological Society of America*, 84, 366–376.

Schelter, B., M. Winterhalder, and J. Timmer, 2006, *Handbook of time series analysis: Recent theoretical developments and applications*: John Wiley & Sons, pp. 514.

Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo, 1998, A comparison of select trigger algorithms for automated global seismic phase and event detection: *Bulletin of the Seismological Society of America*, 88, 95–106.