GEOPHYSICS®

A k-mean characteristic function for optimizing STA/LTA based detection of microseismic events

Journal:	<i>Geophysics</i>
Manuscript ID	GEO-2018-0484.R1
Manuscript Type:	Technical Paper
Keywords:	microseismic, passive seismic, signal processing
Area of Expertise:	Passive Seismic and Microseismic Methods, Signal Processing

SCHOLARONE™
Manuscripts

A *k*-mean characteristic function for optimizing STA/LTA based detection of microseismic events

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

Right Running Head: Optimizing STA/LTA based event detection

*Corresponding author (jubran.akram@kaust.edu.sa)

¹King Abdullah University of Science and Technology (KAUST), Earth Science and Engineering Program, Thuwal, Saudi Arabia. Email: jubran.akram@kaust.edu.sa; daniel.peter@kaust.edu.sa.

²University of Calgary, Department of Geoscience, Calgary, Canada.
Email: eatond@ucalgary.ca.

Abstract

Event detection is an essential component of microseismic data analysis. This process is typically carried out using a short- and long-term average ratio (STA/LTA) method, which is simple and computationally efficient but often yields inconsistent results for noisy datasets. Here, we aim to optimize the performance of the STA/LTA method by testing different input forms of three-component waveform data and different characteristic functions (CFs), including a proposed k -mean CF. These tests are evaluated using receiver operating characteristic (ROC) analysis and compared based on synthetic and field data examples. Our analysis shows that the STA/LTA method using a k -mean CF improves the detection sensitivity and yields more robust event detection on noisy datasets than some previous approaches. In addition, microseismic events are detected efficiently on field data examples using the same detection threshold obtained from the ROC analysis on synthetic data examples. We recommend the use of Youden index based on ROC analysis using a training subset, extracted from the continuous data, to further improve the detection threshold for field microseismic data.

Keywords

Microseismic, event detection, STA/LTA, *k*-nearest neighbors, detection threshold, receiver operating characteristic (ROC) analysis

INTRODUCTION

Event detection is an important processing step that is routinely used in the waveform analysis of both weak (micro) and strong (macro) earthquakes. A rudimentary approach to detect (micro) seismic events involves scanning the waveforms manually. However, this approach is not optimal for large microseismic datasets, because manual scanning of typically complex and noisy waveforms is difficult, time-consuming and prone to subjectivity bias. In contrast, automatic event-detection methods offer a time-efficient and less subjective solution. Numerous automatic event-detection approaches exist in the literature, but the short and long-term average ratio (STA/LTA) method (Allen, 1978; Baer and Kradolfer, 1987; Earle and Shearer, 1994; Withers et al., 1998) is the most widely used, for both strong-motion seismology and in the petroleum industry, to analyze passive seismic data (McEvilly and Majer, 1982; Ruud and Husebye, 1992; Patanè and Ferrari, 1999; Trnkoczy, 2002; Pikoulis and Psarakis, 2010; Schaff, 2010; Qingkai and Ming, 2012; Vaezi and van der Baan, 2015; Wu et al., 2016; Saad et al., 2018).

Like other automatic event-detection methods, the performance of the STA/LTA method typically degrades with decreasing signal-to-noise ratio (S/N) of input waveforms. Careful selection of characteristic function (CF), window lengths, and detection threshold is required in the STA/LTA method for optimal performance (Trnkoczy, 2002; Vaezi and van der Baan, 2015). However, this is quite a challenging task for noisy datasets. For example, considerable effort is required to select an optimal detection threshold in the STA/LTA method. Using a high threshold value in the STA/LTA method results in missing many weak events, whereas a low threshold value yields a higher false positive rate (Sharma, 1991; Liao et al., 2010; Schaff, 2010; Vaezi and van der Baan, 2015). Furthermore, window lengths depend on the typically unknown signal characteristics. Therefore, input datasets are filtered with different frequency pass-bands, and STA/LTA responses

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

are computed using variable size windows (shorter lengths for higher frequencies) to improve event detection (Withers et al., 1998; Baranov, 2007; Schaff, 2010).

In this paper, we aim to improve the performance of the STA/LTA method for event detection on noisy microseismic datasets, by considering the following:

- What should be the S/N of microseismic events for optimal detection by an STA/LTA algorithm? Here, optimal detection implies maximizing the true positive rate while minimizing the false positive rate.
- Can we improve the robustness of the STA/LTA method for noisy datasets by enhancing the performance of the CF?

For this purpose, we review the event-detection workflow using the STA/LTA method for three-component (3C) array data. First, we discuss different input forms that can be generated from the 3C waveform recordings and used in the STA/LTA method. Second, eight different CFs and their corresponding STA/LTA responses are compared to understand their respective detection sensitivity. Moreover, we propose a k -mean CF for achieving robust event detection for noisy datasets. Third, we investigate optimal threshold selection using receiver operating characteristic (ROC) analysis. Finally, we compare the detection results from synthetic and field data examples, using the CF proposed by Allen (1978) and k -mean CF in the STA/LTA method.

Here, we define the key terminology that will be used frequently in the subsequent sections for describing event detection and for evaluating the performance of the STA/LTA method. It includes the following (de Blay et al., 1993; Fluss et al., 2005; Brown and Davis, 2006; Ekelund, 2012):

- **True positives (TP)** are microseismic events that are detected by the STA/LTA (or any other) method as microseismic events.

- **False positives (FP)** are instances of noise that are incorrectly detected as microseismic events.
- **True negatives (TN)** are instances of noise that are correctly identified as noise.
- **False negatives (FN)** are microseismic events that are incorrectly identified as noise.
- **True positive rate (TPR)** is also known as sensitivity. It is the ratio of true positives to the sum of true positives and false negatives

$$TPR = \frac{TP}{TP + FN} . \quad (1)$$

- **True negative rate (TNR)** is also known as specificity. It is the ratio of true negatives to the sum of true negatives and false positives

$$TNR = \frac{TN}{TN + FP} . \quad (2)$$

- **False positive rate (FPR)** is the ratio of false positives to the sum of false positives and true negatives

$$FPR = \frac{FP}{TN + FP} = 1 - TNR . \quad (3)$$

THE SHORT- AND LONG-TERM AVERAGE RATIO (STA/LTA) METHOD

Algorithm

The average of a characteristic function (CF) such as absolute amplitudes or energy of the input signal is computed in two moving windows as follows:

$$(STA)_i = \frac{1}{N_{STA}} \sum_{j=i+1}^{i+N_{STA}} (CF)_j , \quad (4)$$

$$(LTA)_i = \frac{1}{N_{LTA}} \sum_{j=i-N_{LTA}+1}^i (CF)_j , \quad (5)$$

where N_{STA} and N_{LTA} are the STA and LTA window lengths (in terms of samples), respectively. Examples of characteristic functions are given in the following section. The smaller or short-term window provides information on the signal fluctuations whereas larger or long-term window is representative of the background noise level. The ratio of averages in these two windows (STA/LTA), therefore, approximates the local S/N. A detection is triggered when the STA/LTA response exceeds a user-specified threshold value, ϵ (Allen, 1978; Earle and Shearer, 1994; Withers et al., 1998; Sharma et al., 2000; Chen and Stewart, 2005; Diehl et al., 2009; Pikoulis and Psarakis, 2010; Qingkai and Ming, 2012; Vaezi and van der Baan, 2015; Jakka and Garg, 2015; Kalkan, 2016; Velasco et al., 2016), such that

$$\frac{STA}{LTA} \geq \epsilon. \quad (6)$$

Figure 1 shows an event-detection workflow for 3C array data using the STA/LTA method. For downhole microseismic monitoring, arrays comprising up to 50 receiver levels are commonly used (Du and Warpinski, 2013). The detection curves are generated by applying the STA/LTA method (Eqs. 4 to 6) to 3C waveforms at individual receiver levels. Multiple arrivals (similar or different amplitudes) from the same event are typically discarded in the STA/LTA method based on the maximum expected S-P times. Another criterion verifies whether detection is triggered on a user-specified minimum number of receiver levels within a small window. By doing so, inconsistent detections are discarded automatically. Additional checks and quality control are then performed on these detections to identify microseismic events (signal) from false positives (noise).

Two main requirements for improving the performance of the STA/LTA method include the following:

- choosing a characteristic function that minimizes the noisy background fluctuations and improves the signal-to-noise ratio in the analysis window;
- choosing a detection threshold that optimally distinguishes between signal and noise, thus minimizing the false positive rate.

In the subsequent sections, we will review these requirements in detail.

Input: Continuously recorded three-component (3C) waveforms

Because of the large size of recorded microseismic data, input waveforms ($\mathbb{R}^M \times 3$) are typically analyzed in small overlapping windows ($\mathbb{R}^N \times 3$) to avoid memory limitations. These waveforms are mean-corrected to remove any constant voltage offset (DC offset). Moreover, a bandpass filter is applied to suppress unwanted frequencies. From these filtered waveforms, we compute a characteristic function by analyzing each sample individually or a finite sample size window in a sliding fashion (Akram and Eaton, 2016). The STA/LTA response can be computed using any of the following strategies:

- A single representative input can be obtained by stacking the CFs from the 3C waveforms. For example, input CF can represent the stack of absolute amplitudes or envelopes (Oye and Roth, 2003) or the vector magnitudes of 3C waveforms, $\sqrt{\sum_{i=1}^3 x_i^2}$ (Bai and Kennett, 2000; Baranov, 2007; Ross and Ben-Zion, 2014). An STA/LTA response is computed for the input CF and a threshold is applied to detect microseismic events.
- For each component, we can compute the CF and the corresponding STA/LTA response. Microseismic events can be detected either on individual components or the stacked STA/LTA response; however, the former requires an additional criterion which depends on the number of triggers on any or as many components (Velasco et al., 2016).

Figure 2 shows the STA/LTA responses computed using the strategies mentioned above. In these examples, we arbitrarily select the absolute amplitude of the input waveform as the CF to compute the STA/LTA response, which for both strategies are almost identical. Henceforth, we will stack the CFs from each component to compute the input to be used in the STA/LTA method.

Characteristic functions

Numerous CFs have been used as input in the STA/LTA method, including the absolute amplitude, root-mean-square (RMS) amplitude, energy, maximum energy, or the mean energy of a microseismic trace and its derivatives. These CFs are computed either at each sample individually or in a finite size (related to the dominant period of the signal) sliding window (Allen, 1978; McEvilly and Majer, 1982; Earle and Shearer, 1994; Bai and Kennett, 2000; Oye and Roth, 2003). Here, we propose a novel CF which is inspired by the k nearest neighbors (KNN) method from machine learning. The KNN method has been used previously to solve many classification and regression problems (Dunham, 2003; Schelter et al., 2006; Lindenbaum et al., 2016). In seismic detection problems, the KNN method is typically applied to the STA/LTA response for classifying signal and noise intervals (e.g. Lindenbaum et al., 2016). However, we compute the distance of each sample (in terms of amplitude) from other samples within an analysis window and use the mean value of k nearest neighbors to generate a CF, which we hereby call k -mean CF. This CF yields better minimization of the background fluctuations as compared to the amplitude- or energy-based CFs that are directly obtained from each sample's amplitude and do not use information from the surrounding samples. The workflow to compute k -mean CF is described as follows:

Step 1: For the p th component, compute the Euclidean distance (d) of amplitude (x) at each sample from the remaining samples

$$d_{ij}^{(p)} = |x_i^{(p)} - x_j^{(p)}|, \quad p = 1, 2, 3 \text{ and } i, j = 1, 2, 3, \dots, N_{\text{samples}}, \quad (7)$$

where $N_{samples}$ is the total number of samples in the analysis window.

Step 2: Sort each row of $d_{ij}^{(p)}$ in ascending order.

Step 3: Compute the mean value of the first k values in each row of the sorted matrix $\tilde{d}_{ij}^{(p)}$

$$K_i^{(p)} = \frac{1}{k} \sum_{j=1}^k \tilde{d}_{ij}^{(p)}. \quad (8)$$

Step 4: Stack K from each of the three components to generate k -mean CF.

Figure 3 shows a comparison of five different CFs and their corresponding STA/LTA responses. For a relatively low S/N input waveform that contains both P- and S-wave arrivals, we compute absolute amplitude, RMS amplitude, energy, the CF described in Allen (1978) and k -mean CF. Allen's CF is the weighted sum of energy of a microseismic trace and its derivative. The CFs including absolute amplitude, energy and Allen's CF directly use the amplitude of each sample and do not incorporate information from the surrounding samples. Although CFs such as RMS amplitude are obtained in a moving window fashion, these fail to minimize the background fluctuations as efficiently as the k -mean CF. Consequently, these CFs do not improve the S/N as much as the k -mean CF which uses a distance-based metric within the analysis window. Because the majority of the samples in the noise interval have similar amplitudes, taking the mean of first k smallest distance values for each sample will result in a much lower and stable background noise level. The analysis window is typically designed as such to contain only a single event at a time apart from events that are recorded at the almost similar time. In the signal interval, the distance of each sample from other samples within the analysis window is much greater which results in a higher k -mean value. The STA/LTA responses for these CFs are computed using three different STA window lengths $((1, 2, 3) \times \tau_D)$. The LTA window length is chosen as 5 times the STA

1
2
3 window length (Akram and Eaton, 2016). The STA/LTA response at the signal interval is the
4
5 highest when the STA window length = τ_D . However, the STA/LTA response is more sensitive to
6
7 noise fluctuations, which is not desired in the event detection as it makes the threshold selection
8
9 difficult. To minimize the background fluctuations, we recommend choosing the STA window
10
11 length between $2\tau_D$ and $3\tau_D$. We discuss the window length in more detail in the next section. In
12
13 all cases, the STA/LTA response for the k -mean CF outperforms the STA/LTA responses for the
14
15 rest of CFs, mainly due to higher stability in the background noise level and higher S/N of the k -
16
17 mean CF.

18
19 Figure 4a-c shows another comparison of eight different CFs including absolute amplitude,
20
21 envelope, maximum amplitude, RMS amplitude, energy, maximum energy, and Allen's CF. All
22
23 amplitude-based CFs yield similar STA/LTA responses, while all energy-based CFs also yield
24
25 similar STA/LTA responses. Although the STA/LTA responses from all CFs exhibits distinct
26
27 fluctuations at the signal intervals, the k -mean CF yields the strongest response for the weakest
28
29 event, in this case, seen at 2s (Figure 4d). To quantify the comparisons, we define the detection
30
31 sensitivity as the ratio of peak STA/LTA response in the signal interval containing the weakest
32
33 event to the peak STA/LTA response in the noise interval in the analysis window. A lower
34
35 detection sensitivity implies that a small change in the detection threshold value can significantly
36
37 affect the number of triggers as well as the false positive rate. By comparing the detection
38
39 sensitivity (Figure 4e) for all CFs, we found that amplitude-based CFs have lower detection
40
41 sensitivity as compared to the energy-based CFs. Our proposed k -mean CF yields the highest
42
43 detection sensitivity and, therefore, is superior to amplitude-based CFs for noisy datasets to
44
45 achieve optimal detection.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Window lengths

The STA and LTA window lengths control the degree to which a signal can be distinguished from the background on the STA/LTA response. An STA window that is too short can produce meaningless noise fluctuations on the STA/LTA response, thus yielding a higher false positive rate. Conversely, a very long LTA window will smooth the background noise too aggressively, and the corresponding STA/LTA response could miss closely spaced events. Numerous authors have recommended that STA and LTA window lengths should be based on the frequency of the expected signal in the recorded waveforms (e.g., Earle and Shearer, 1994; Tong, 1995; Withers et al., 1998; Akram and Eaton, 2016). An STA window length is typically chosen to be 2-3 times the dominant signal period (τ_D) whereas an LTA window length is 5-10 times longer than the STA window length (Tong, 1995; Akram and Eaton, 2016).

Figure 5 shows the peak STA/LTA response in the signal interval for both microseismic events (at 1s and 2s in Figure 4). In Figures 5a and 5c, STA/LTA response was computed for STA window lengths ranging between $1.5\tau_D$ and $5\tau_D$ (in this case, $\tau_D = 21$ samples = 0.0105s) and the LTA window lengths equal to 8 times the STA window length. In Figures 5b and 5d, STA/LTA response was computed by using STA window length = $2.5\tau_D$ and a range of LTA window lengths (4 – 10 times the STA window length). The STA/LTA response is higher when a smaller STA window is used. However, the STA/LTA response shows meaningless fluctuations when the STA window length is very small. Also, the STA/LTA response is higher when a longer LTA window is used. However, the STA/LTA response fails to resolve closely spaced events when the LTA window length is very large. Therefore, optimal STA window length lies between $2\tau_D$ and $3\tau_D$ whereas optimal LTA window length lies between 5 and 10 times the STA window length. The STA/LTA values for both Allen's and k -mean CFs are higher than 10 for the selected ranges of STA and

LTA window lengths. By using a suitable threshold value ($\epsilon > 3$), we can easily detect both microseismic events. However, for the weaker microseismic event, the proposed k -mean CF yields higher STA/LTA values and is, therefore, more robust for noisy datasets.

Detection threshold

A threshold criterion is required for detection of microseismic events. The static threshold comprises a user-specified fixed value whereas a dynamic or adaptive threshold is data dependent and time-varying (e.g., Xiantai et al., 2011). For noisy datasets, the selection of an optimal static threshold is often challenging. The STA/LTA response for the signal corresponding to a weaker microseismic event can be very similar to the background noise level. Consequently, using a higher threshold value may result in missing weaker events whereas a slightly lower threshold may yield a higher false positive rate. Because the STA/LTA responses from energy-based CFs and the k -mean CF have higher detection sensitivity (Figure 4e), using these CFs for noisy datasets can, therefore, minimize the false positive rate.

We can further improve the performance of the STA/LTA method by computing an optimal threshold value from the receiver operating characteristic (ROC) analysis (de Blay et al., 1993; Fluss et al., 2005; Ekelund, 2012). The ROC curve gives a relationship between FPR and TPR for every possible cut-off or detection threshold value. A detection threshold is considered optimal if it yields the lowest FPR and the highest TPR possible. In the ROC analysis, this optimal threshold value can be obtained using Youden index (J) which corresponds to the detection threshold that maximizes $TPR - FPR$, such that $J(\epsilon) = \max_{\epsilon} \{TPR(\epsilon) - FPR(\epsilon)\}$ (Unal, 2017).

APPLICATION ON A SYNTHETIC MICROSEISMIC DATASET

We generated synthetic waveforms for 100 microseismic events by using the far-field displacement equations (Lay and Wallace, 1995; Aki and Richards, 2002; Eaton, 2018)

$$u_i^P(x, t) = \frac{\gamma_i \gamma_j \gamma_k M_{jk} \dot{w}(t - \tau_P)}{4\pi r \rho \alpha^3}, \quad (9)$$

$$u_i^S(x, t) = \frac{(\delta_{ij} - \gamma_i \gamma_j) \gamma_k M_{jk} \dot{w}(t - \tau_S)}{4\pi r \rho \beta^3}, \quad (10)$$

where $w(t)$ is the displacement time function at the source, and dot notation represents a time derivative; γ_i is the direction cosine of the ray at the source location with the coordinate axis x_i ; r is the source-receiver distance; τ_P is the traveltime from source to the receiver; and, ρ , α , and β are the density and the velocities of P- and S-wave at the source, respectively. In addition, M_{jk} represents a moment-tensor component comprised of a force couple that is composed of two equal amplitude, but opposite forces in the $\pm j$ direction, separated by an infinitesimal distance in the k direction. The seismic moment tensor M is written as

$$M = M_0 \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}, \quad (11)$$

where M_0 is the scalar seismic moment in Nm, which is related to the moment magnitude as (Hanks and Kanamori, 1979; Daniel, 2014)

$$M_W = \frac{2}{3}(\log_{10} M_0 - 9.1). \quad (12)$$

A summation convention is used in Eqs. 9 and 10 for repeated indices. The microseismic data were recorded with a vertical array of 12 3C receivers (Figure 6). The receiver interval and the sampling

interval to record these waveforms were 15m and 0.5ms, respectively. For the traveltime and displacement computations, density, P-wave velocity, and S-wave velocity values were 2200kg/m³, 5000m/s, and 2941m/s, respectively. The moment magnitudes were uniformly distributed between Mw = -3 and Mw = -1 over all microseismic events. Also, a single source mechanism (moment tensor: $M = [M_{11} \ M_{22} \ M_{33} \ M_{12} \ M_{13} \ M_{23}]^T = [0 \ 0 \ 0 \ 1 \ 0 \ 0]^T$) was used to generate synthetic waveforms.

To mimic a real experiment, we generated a uniform Gaussian noise series and embedded synthetic microseismic events within to form a continuously recorded waveform segment. Figure 7 shows a 7.5s long subinterval that contains seven microseismic events of different magnitudes. Because we used a constant background noise level, the magnitude level directly reflects the S/N of recorded waveforms, which represents the ratio of maximum amplitude (P or S) in the signal interval to the maximum noise level in the analysis window. For a source gather, the mean value of S/N of all receivers within the array is used. The final S/N is then the average S/N of all three components. For stronger events (Mw = -2.5 and above), both P- and S-wave arrivals are visible on all receiver levels. However, the computed S/N is less than one for lower magnitude events (Mw = -2.75 and below). The P-wave arrival is barely visible on most of the receivers for the event at 5s (Mw = -2.75). Although the S-wave arrival can easily be seen on all components, its maximum value on most receivers in x₁- and x₂-components is close to the maximum noise level and, therefore, averaging S/N across all components and receivers yields an S/N value which is less than 1 (0.99).

Figure 8a shows the change in the number of triggers with the detection threshold value. The number of triggers is determined from the STA/LTA responses obtained using Allen's and *k*-mean CFs. For the STA/LTA computation, the STA window length was 2.5τ_D whereas LTA window length was 7 times longer than the STA window length.

The step size for the detection threshold in both cases is the same ($\Delta\epsilon = 0.2$). In case of STA/LTA response using Allen's CF, the number of triggers increases rapidly to around 350 (in two steps) when the threshold value is decreased from the optimal value ($\epsilon = 1.95$). However, the increase in the number of triggers is not as rapid for STA/LTA response using k -mean CF, thus suggesting that the proposed CF is more robust in noisy datasets. Figure 8b shows the TPR and FPR for both cases. Our goal is to find a threshold value that provides the highest possible TPR with the lowest FPR (the top left corner of the curve). This optimal cut-off value is obtained using the Youden index, which is the detection threshold value corresponding to the maximum of TPR-FPR (Figure 8c). For the synthetic data example, the optimal thresholds (Youden index) for Allen's CF and k -mean CF are $\epsilon = 1.95$ and $\epsilon = 3.05$, respectively.

Figure 9a-e shows the S/N and Mw values and the STA/LTA detection curves for the complete synthetic microseismic data (100 events). Around 10% of the generated microseismic events have $S/N \leq 1$ (Figure 9a). The corresponding magnitudes of these events are also less than $M_w = 2.75$ (Figure 9b). By using the optimal threshold value ($\epsilon = 1.95$) on the STA/LTA response obtained using Allen's CF, we were able to detect 98 of 100 microseismic events (Figure 9c). Both missed events have $S/N < 1$. Figure 9d shows the STA/LTA detection curve for the k -mean CF with an optimal threshold value ($\epsilon = 3.05$). In this case, we were able to detect 97 out of 100 microseismic events, where all missed events have $S/N < 1$. Although STA/LTA method using Allen's CF yields one more detection, the S/N of the detected event is less than 1 and, therefore, cannot always be reliably detected by the STA/LTA method. For a higher threshold value ($\epsilon = 4.0$), the STA/LTA method using Allen's CF could not detect 10 microseismic events (Figure 9e) whereas only 4 microseismic events were missed when k -mean CF is used (Figure 9f). As can be seen in Figure 8a, the STA/LTA method using Allen's CF yields identical results for threshold

value ($\epsilon = 3.0$). This result suggests that the detection sensitivity is improved when k -mean CF is used in the STA/LTA method. Also, small changes in the threshold value do not significantly affect the detection results in the proposed CF case.

APPLICATION ON A REAL MICROSEISMIC DATASET

Real microseismic data examples were acquired by a borehole receiver array in western Canada. For this study, we received only a finite number of previously detected microseismic events (both low and high S/N). To mimic a continuously recorded waveform segment, we combined several microseismic events, which were arranged in a way such that a high S/N event follows a low S/N. The intent of this is to test the STA/LTA detector on a dataset with a lot of S/N variation and closely spaced events (a more challenging scenario). Figure 10 shows the 3C waveforms that contain five microseismic events, among which three events are low S/N and the remaining two are relatively high S/N. For the stronger events, both P- and S-wave arrivals are visible on all three components.

Figure 11 shows the event detection on receiver level 10 from the real microseismic data example. Because the signal characteristics are typically unknown, a bank of bandpass filters (20-100Hz, 50-200Hz, 150-300Hz) were applied to the input data. Figure 11a shows the raw 3C waveforms and the corresponding bandpass filtered responses. The dominant signal periods were computed for each band and used to determine the STA and LTA window lengths. Figure 11b shows the STA/LTA responses when Allen's CF is used. The optimal threshold value ($\epsilon = 1.95$) estimated from the synthetic example in Figure 8 was first applied on the real microseismic data. By using this optimal threshold, we were able to detect all five events on the raw waveforms. However, for other filtered cases, the optimal threshold value yielded false positives, and it was considered too low for 150-300Hz filtered waveforms. A higher threshold value ($\epsilon = 3.05$) failed to detect two

events (at 46.5s and 48.5s) for the raw waveforms. Similarly, it missed the event at 44.5s on 150-300Hz waveforms. However, all five events were detected without false positives in 20-100Hz and 50-200Hz filtered waveforms. In the case of STA/LTA response with k -mean CF (Figure 11c), the optimal threshold value ($\epsilon = 3.05$) from Figure 7 yields all five events for the raw and filtered waveforms. The background of the STA/LTA response also does not fluctuate as in the case of Allen's CF. For example, the major difference can be observed in the STA/LTA response for the 150-350Hz filtered waveforms. Due to the higher detection sensitivity of STA/LTA response for the k -mean CF, the optimal threshold value from the synthetic example also works better on real data. In practice, we can perform the ROC analysis on a training dataset (a subset of continuously-recorded real microseismic data) to find the optimal detection threshold.

DISCUSSION

Impact of k value on STA/LTA

To understand the impact of k value for the k -mean CF on the STA/LTA response, we compute the peak STA/LTA response for six events ($S/N > 1$) from the synthetic example (Figure 7). Figures 12a-f show the peak STA/LTA response against the k -value that ranges between $0.5\tau_D - 3\tau_D$. We notice that increasing the k -value decreases the peak STA/LTA response. However, for all events, peak STA/LTA response decreases rapidly after $k = 1.5\tau_D$. For smaller k values (less than $0.5\tau_D$), the k -mean CF becomes noisy and therefore, corresponding STA/LTA response will yield more false positives. Higher k values will also decrease the STA/LTA response and, therefore, the detection in noisy datasets can become challenging. Thus, for optimal event detection, k -values should be chosen between $1\tau_D - 2\tau_D$.

S/N limit for optimal STA/LTA based detection

According to Schaff (2010), an STA/LTA method is incapable of detecting signals that are buried in the background noise level ($S/N \leq 1$). Considering the inherent design of the STA/LTA method, it is difficult to detect events that are characterized by waveforms with $S/N \leq 1$. We have shown in Figure 9 that the STA/LTA method becomes unstable as the S/N of input waveforms approaches one. However, this is often the case for single receiver levels. In practice, microseismic data are typically recorded with one or more arrays comprising up to 50 receiver levels in wells (Du and Warpinski 2013). In many cases, P- and S-wave arrivals from a microseismic event are visible only on some receivers as we have seen in Figure 7. Unless a microseismic event is invisible on all receiver levels, it can be detected by the STA/LTA method using a suitable minimum receiver level criterion. We have shown using synthetic data examples that it is possible to detect weaker microseismic events (S/N closer to but higher than 1) optimally with an STA/LTA method. Because the STA/LTA responses for weaker events suffer from low detection sensitivity (response to a threshold value), the STA/LTA method yields a higher false positive rate for noisy datasets as seen in Figure 7. Using a suitable characteristic function, we can improve the detection sensitivity of the STA/LTA method and minimize the false positive rate. We have shown in Figures 3 and 4 that the amplitude-based CFs have lower detection sensitivity and are, therefore, not suitable for event detection on noisy datasets. For noisy datasets, selection of a detection threshold is often challenging because of the small differences between the STA/LTA values in the noise and signal intervals. The STA/LTA method using energy-based CFs and k -mean CF yields higher detection sensitivity. Besides, the optimal detection threshold for the noisy dataset can be obtained from the Youden index in the ROC analysis as seen in Figures 8 and 9. The STA/LTA method, therefore,

becomes more robust for event detection on noisy datasets when k -mean CF is used, and the optimal detection threshold is selected from the ROC analysis.

CONCLUSIONS

This paper provides a review of the conventional STA/LTA method to detect microseismic events and describe various input forms obtained from 3C waveforms for use in the STA/LTA method. We find that the STA/LTA response obtained from stacking the CFs for the three receiver components is almost identical to the stack of STA/LTA responses from individual components. Eight different characteristic functions and their corresponding STA/LTA responses are investigated to understand the detection sensitivity, showing that amplitude-based CFs produce lower detection sensitivity than the energy-based CFs. We propose a novel k -mean CF that yields the highest detection sensitivity in the STA/LTA method among the considered CFs. The receiver operating characteristic (ROC) analysis is a useful technique for finding optimal thresholds and should be used on a training subset from the continuously recorded real microseismic data for improved detection results. For the proposed k -mean CF, we find that optimal threshold values obtained from ROC analysis on synthetic data examples also work well for real data examples. The event detection results based on synthetic data examples show that it is possible to detect weaker events optimally with an STA/LTA algorithm; however, the number of false positives increases significantly when weaker events are desired. The STA/LTA method becomes unstable as the S/N of input waveforms approaches one. For noisy datasets, selection of a detection threshold is often challenging because of the small differences between the STA/LTA values in the noise and signal intervals (detection sensitivity of the STA/LTA method is low). To increase the robustness of the STA/LTA method, we recommend using energy-based CFs and k -mean CF for higher detection sensitivities and the Youden index in the ROC analysis for optimal detection

threshold. Using our proposed k -mean CF in the STA/LTA method yields more robust event detection on noisy datasets as compared to the characteristic functions based on the amplitude and energy of the input waveforms.

REFERENCES

Aki, K., and P. G. Richards, 2002, Quantitative Seismology (2nd Edition): *University Science Books*, pp. 700.

Akram, J., and D. W. Eaton, 2016, A Review and Appraisal of Arrival-Time Picking Methods for Downhole Microseismic Data: *Geophysics*, 80 (2), KS67-KS87.

Allen, R., 1978, Automatic Earthquake Recognition and Timing from Single Traces: *Bulletin of the Seismological Society of America*, 68, 1521–1532.

Baer, M., and U. Kradolfer, 1987, An Automatic Phase Picker for Local and Teleseismic Events: *Bulletin of the Seismological Society of America*, 77, 1437-1445.

Bai, C., and B. L. N. Kennett, 2000. Automatic Phase-Detection and Identification by Full Use of a Single Three-Component Broadband Seismogram: *Bulletin of the Seismological Society of America*, 90 (1), 187–198.

Baranov, S. V., 2007, Application of the Wavelet Transform to Automatic Seismic Signal Detection: *Physics of the Solid Earth*, 43 (2), 177-188.

Brown, C. D., and H. T. Davis, 2006, Receiver operating characteristics curves and related decision measures: A tutorial: *Chemometrics and Intelligent Laboratory Systems*, 80, 24 – 38.

- Caffagni, E., D. W. Eaton, J. P. Jones, and M. van der Baan, 2016, Detection and Analysis of Microseismic Events using a Matched Filtering Algorithm (MFA): *Geophysical Journal International*, 206, 644–658.
- Chen, Z., and R. Stewart, 2005, Multi-window Algorithm for Detecting Seismic First Arrivals: *CSEG National Convention*, 1-4.
- Daniel, G., 2014, Bias in magnitude for earthquakes with unknown focal mechanism: *Geophysical Prospecting*, 62, 848–861.
- de Blay, F., H. Zana, M. Offner, A. Verot, M. Velten, and G. Pauli, 1993, Receiver Operating Characteristic Analysis: A Useful Method for a Comparison of the Clinical Relevance of Two in Vitro IgE Tests: *Journal of Allergy and Clinical Immunology*, 92 (2), 255-263.
- Diehl, T., N. Deichmann, E. Kissling, and S. Husen, 2009, Automatic S-wave picker for local earthquake tomography: *Bulletin of the Seismological Society of America*, 99, 1906–1920.
- Du, J., and N. R. Warpinski, 2013, Velocity Building for Microseismic Hydraulic Fracture Mapping in Isotropic and Anisotropic Media: *SPE 163866*.
- Dunham, M. H., 2003, Data Mining: Introductory and Advanced Topics: *Prentice Hall, New Jersey*, pp. 313.
- 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.
- Eaton, D.W., 2018, Passive Seismic Monitoring of Induced Seismicity: Fundamental Principles and Application to Energy Technologies: Cambridge University Press.

Ekelund, S., 2012, ROC Curves – What are They and How are They Used? *Point of Care*, 11 (1), 16-21.

Fluss, R., D. Faraggi, and B. Reiser, 2005, Estimation of the Youden Index and its Associated Cutoff Point: *Biometrical Journal*, 47 (4), 458-472.

Hanks T. C., and H. Kanamori, 1979, A moment magnitude scale: *Journal of Geophysical Research*, 84 (B5), 2348–2350.

Jakka, R. S., and S. Garg, 2015, Suitable Triggering Algorithms for Detecting Strong Ground Motions Using MEMS Accelerometers: *Earthquake Engineering and Engineering Vibration*, 14 (1), 27-35.

Kalkan, E., 2016, An Automatic P-Phase Arrival-Time Picker: *Bulletin of the Seismological Society of America*, 106 (3), 971–986.

Lay, T., and T. C. Wallace, 1995, Modern Global Seismology: *Academic Press*, pp. 521.

Liao, B., C. Chen, and T. Sheu, 2010, Using a Fractal Analysis and Polarization Method for Phase Identification in Three-Component Seismograms: *Terr. Atmos. Ocean. Sci.*, 21 (5), 791-805.

Lindenbaum, O., N. Rabin, and Y. Bregnan, 2016, Multi-channel Fusion for Seismic Event Detection and Classification: *International Conference on the Science and Engineering (ICSEE)*, pp. 5.

McEvelly, T. V., and E. L. Majer, 1982, ASP: An Automated Seismic Processor for Microearthquake Networks: *Bulletin of the Seismological Society of America*, 72 (1), 303–325.

- Oye, V., and M. Roth, 2003, Automated Seismic Event Location for Hydrocarbon Reservoirs: *Computers & Geosciences*, 29, 851–863.
- Patanè, D., and F. Ferrari, 1999, ASDP: a PC-based Program using a Multi-algorithm Approach for Automatic Detection and Location of Local Earthquakes: *Physics of the Earth and Planetary Interiors*, 113, 57-74.
- Pikoulis, E. M., and E. Z. Psarakis, 2010, A Ratio Test for the Accurate Automatic P-wave Onset Detection: *ICSP Proceedings*, 2621-2624.
- Qingkai, K., and Z. Ming, 2012. Evaluation of Earthquake Signal Characteristics for Early Warning: *Earthquake Engineering and Engineering Vibration*, 11 (3), 435–443.
- Ross, Z. E., and Y. Ben-Zion, 2014. An Earthquake Detection Algorithm with Pseudo-Probabilities of Multiple Indicators: *Geophysical Journal International*, 197, 458–463.
- Ruud, B. O., and E. S. Husebye, 1992, A New Three-Component Detector and Automatic Single-Station Bulletin Production: *Bulletin of the Seismological Society of America*, 82 (1), 221–237.
- Saad, O. M., A. Shalaby, L. Samy, and M. S. Sayed, 2018, Automatic Arrival Time Detection for Earthquakes based on Modified Laplacian of Gaussian filter: *Computers & Geosciences*, 113, 43-53.
- Schaff, D., 2010. Improvements to Detection Capability by Cross-correlating for Similar Events: a Case Study of the 1999 Xiuyan, China, Sequence and Synthetic Sensitivity Tests: *Geophysical Journal International*, 180, 829–846.

Schelter, B., M. Winterhalder, and J. Timmer, 2006, Handbook of Time Series Analysis: Recent Theoretical Developments and Applications: John Wiley & Sons, pp. 514.

Sharma, B. K., A. Kumar, and V. M. Murthy, 2010. Evaluation of seismic event-detection algorithms: *Journal Geological Society of India*, 75, 533–538.

Song, F., H. S. Kuleli, M. N. Toksöz, E. Ay, and H. Zhang, 2010. An improved method for hydrofracture-induced microseismic event detection and phase picking: *Geophysics*, 75 (6), A47–A52.

Tan, Y., C. He, J. Yu, and G. Feng, 2014. A combined method for automatic microseismic event detection and arrival picking: *84th Annual International Meeting, SEG, Expanded Abstracts*, 2335–2340.

Tong, C., 1995. Characterization of Seismic Phases - An Automatic Analyser for Seismograms: *Geophysical Journal International*, 123, 937–947.

Trnkoczy, A., 2002, Understanding and Parameter Setting of STA/LTA Trigger Algorithm, in IASPEI New Manual of Seismological Observatory Practice, 2, pp. 01–20, ed. Bormann, P.

Unal, I., 2017, Defining an Optimal Cut-Point Value in ROC Analysis: An Alternative Approach: *Computational and Mathematical Methods in Medicine*, Article ID 3762651, pp. 14.

Vaezi, Y., and M. van der Baan, 2015. Comparison of the STA/LTA and power spectral density methods for microseismic event detection: *Geophysical Journal International*, 203, 1896–1908.

- Velasco, A. A., R. Alfaro-Diaz, D. Kilb, and K. L. Pankow, 2016. A Time-Domain Detection Approach to Identify Small Earthquakes within the Continental United States Recorded by the US Array And Regional Networks: *Bulletin of the Seismological Society of America*, 106 (2), 512–525.
- 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.
- Wu, S., Y. Wang, Y. Zhan, and X. Chang, 2016, Automatic Microseismic Event Detection by Band-Limited Phase-Only Correlation: *Physics of the Earth and Planetary Interiors*, 261, 3-16.
- Xiantai, G., L. Zhimin, Q. Na, and J. Weidong, 2011, Adaptive Picking of Microseismic Event Arrival using a Power Spectrum Envelope: *Computers & Geosciences*, 37, 158-164.

List of Figures

Figure 1. Workflow of an STA/LTA detector for 3C array data.....30

Figure 2. STA/LTA responses from 3C waveforms. a) x_1 -, x_2 -, and x_3 -components of input waveforms. Two microseismic events exist at around 1s (strong) and 2s (weak). b) Absolute amplitudes computed from the input waveforms. c) STA/LTA responses for the CFs. d) CF obtained by stacking the absolute amplitudes of each component in b). e) STA/LTA (black) obtained by stacking the responses in c). STA/LTA (red) computed from the CF in d).....31

Figure 3: An example of five different characteristic functions (left) and their corresponding STA/LTA responses (right). These CFs include absolute amplitude, RMS amplitude, energy, Allen’s CF and k -mean CF. The onset of P and S-wave arrivals are indicated by solid vertical lines (magenta). The dashed vertical lines indicate multiples of the dominant period of signal (1= blue, 2 = green, 3 = red) that are used as STA window lengths in this example. The LTA window length in all cases is chosen as 5 times the STA window length.....32

Figure 4. Comparison of STA/LTA responses obtained by using numerous CFs. a) 3C input waveform (same as in Figure 2). b) STA/LTA responses for various amplitude-based CFs. c) STA/LTA responses for various energy-based CFs. d) STA/LTA response for the k -mean curve as CF. The signal interval for both strong and weak microseismic events are shown by a rectangle (dashed red). e) The detection sensitivity of each STA/LTA response (a – d).33

Figure 5. Effect of window size on STA/LTA for both Allen’s and k -mean CFs. STA/LTA is computed for a) weaker event (around 2s in Figure 3) and using STA window lengths ranging between $1.5\tau_D$ and $5\tau_D$ and LTA window length = 8 times the STA window length. b) weaker event and using STA window length = $2.5\tau_D$ and LTA window lengths ranging between 4 and 10

times the STA window length. c) stronger event (around 1s in Figure 3) and using STA window lengths ranging between $1.5\tau_D$ and $5\tau_D$ and LTA window length = 8 times the STA window length. d) stronger event and using STA window length = $2.5\tau_D$ and LTA window lengths ranging between 4 and 10 times the STA window length. The optimal window lengths are indicated by the rectangular boxes (dashed red).34

Figure 6. Source-receiver geometry for synthetic microseismic data. A total of 100 microseismic events were recorded on a vertical array of twelve 3C receivers. All microseismic events shared the same moment tensor ($[0 \ 0 \ 0 \ 1 \ 0 \ 0]^T$) but their magnitudes were uniformly distributed between $M_w = -3$ and $M_w = -1$35

Figure 7. A 7.5s subinterval from the continuous record (with noise added) that includes seven microseismic events with different magnitudes. Large amplitudes associated with stronger microseismic events are clipped for display purpose. A zoomed view around 2s and 5s shows the waveforms for -2.36 and -2.75 magnitude events, respectively. For the latter event, the P wave is not visible on all three components whereas the S-wave amplitude level is almost the same as the background noise.36

Figure 8. Event detection (STA/LTA used Allen's CF and k -mean CF) on the synthetic dataset containing 100 microseismic events. a) Detection threshold versus the number of triggers plot, showing that the smaller the threshold value, the higher the number of triggers. b) False positive rate (FPR) versus true positive rate (TPR) plot. c) Detection threshold versus TPR- FPR. The maximum value of which represents the Youden index and is considered the optimal detection threshold value ($\epsilon = 1.95$ for Allen's CF and $\epsilon = 3.05$ for k -mean CF) on this dataset.37

Figure 9. a) S/N of synthetic microseismic events. Higher S/N values are clipped to zoom in on the events with lower S/N. b) Moment magnitudes (M_w) that are uniformly distributed between -

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 and -3. Detection curve (green) and 3C waveforms for c) optimal threshold value ($\epsilon \equiv 1.95$) from Figure 8 for Allen’s CF. Vertical arrows indicate the missing events. d) optimal threshold value (3.05) from Figure 8 for the k -mean CF. e) threshold value ($\epsilon \equiv 4.0$) for Allen’s CF. f) threshold value ($\epsilon \equiv 4.0$) for the k -mean CF.....38

Figure 10. Example of real data containing five microseismic events in which three have relatively low S/N waveforms.39

Figure 11. Example of event detection on real microseismic data. a) 3C waveform at receiver level 10 from the gather shown in Figure 10. Various bandpass filters (20-100Hz, 50-200Hz, 150-300Hz) were applied to data for extracting useful signal. b) STA/LTA responses computed using Allen’s CF for the 3C input waveform in a). Horizontal lines represent the threshold value used to detect microseismic events (red = optimal threshold ($\epsilon \equiv 1.95$) from synthetic examples, blue = threshold ($\epsilon \equiv 3.05$). The rectangles are used to highlight areas of poor detections. The arrow shows that the optimal threshold value is too low for 150-300Hz filtered waveforms. c) STA/LTA responses computed using k -mean CF. Horizontal line indicates the optimal threshold value ($\epsilon \equiv 3.05$) from synthetic examples in Figure 8. In this case, all five events are detected in all passbands and the number of false positives is lower in 150-300Hz filtered waveforms. The rectangles highlight the similar areas as in b) for comparisons.....40

Figure 12. Plots of k versus peak STA/LTA response for microseismic events from the synthetic data example in Figure 7 at a) 1s, b) 2s, c) 3s, d) 4s, e) 6s, and f) 7s. Vertical lines indicate the signal’s dominant period (dashed) and $1.5\tau_D$ (bold).41

Figure 1. Workflow of an STA/LTA detector for 3C array data.

195x100mm (300 x 300 DPI)

Figure 2. STA/LTA responses from 3C waveforms. a) x_1 -, x_2 -, and x_3 -components of input waveforms. Two microseismic events exist at around 1s (strong) and 2s (weak). b) Absolute amplitudes computed from the input waveforms. c) STA/LTA responses for the CFs. d) CF obtained by stacking the absolute amplitudes of each component in b). e) STA/LTA (black) obtained by stacking the responses in c). STA/LTA (red) computed from the CF in d).

211x141mm (300 x 300 DPI)

Figure 3: An example of five different characteristic functions (left) and their corresponding STA/LTA responses (right). These CFs include absolute amplitude, RMS amplitude, energy, Allen’s CF and k-mean CF. The onset of P and S-wave arrivals are indicated by solid vertical lines (magenta). The dashed vertical lines indicate multiples of the dominant period of signal (1= blue, 2 = green, 3 = red) that are used as STA window lengths in this example. The LTA window length in all cases is chosen as 5 times the STA window length.

207x160mm (300 x 300 DPI)

Figure 4. Comparison of STA/LTA responses obtained by using numerous CFs. a) 3C input waveform (same as in Figure 2). b) STA/LTA responses for various amplitude-based CFs. c) STA/LTA responses for various energy-based CFs. d) STA/LTA response for the k-mean curve as CF. The signal interval for both strong and weak microseismic events are shown by a rectangle (dashed red). e) The detection sensitivity of each STA/LTA response (a – d).

158x169mm (300 x 300 DPI)

Figure 5. Effect of window size on STA/LTA for both Allen's and k-mean CFs. STA/LTA is computed for a) weaker event (around 2s in Figure 3) and using STA window lengths ranging between $1.5\tau_D$ and $5\tau_D$ and LTA window length = 8 times the STA window length. b) weaker event and using STA window length = $2.5\tau_D$ and LTA window lengths ranging between 4 and 10 times the STA window length. c) stronger event (around 1s in Figure 3) and using STA window lengths ranging between $1.5\tau_D$ and $5\tau_D$ and LTA window length = 8 times the STA window length. d) stronger event and using STA window length = $2.5\tau_D$ and LTA window lengths ranging between 4 and 10 times the STA window length. The optimal window lengths are indicated by the rectangular boxes (dashed red).

183x159mm (300 x 300 DPI)

Figure 6. Source-receiver geometry for synthetic microseismic data. A total of 100 microseismic events were recorded on a vertical array of twelve 3C receivers. All microseismic events shared the same moment tensor $([0 \ 0 \ 0 \ 1 \ 0 \ 0]T)$ but their magnitudes were uniformly distributed between $M_w = -3$ and $M_w = -1$.

176x132mm (300 x 300 DPI)

Figure 7. A 7.5s subinterval from the continuous record (with noise added) that includes seven microseismic events with different magnitudes. Large amplitudes associated with stronger microseismic events are clipped for display purpose. A zoomed view around 2s and 5s shows the waveforms for -2.36 and -2.75 magnitude events, respectively. For the latter event, the P wave is not visible on all three components whereas the S-wave amplitude level is almost the same as the background noise.

211x242mm (300 x 300 DPI)

Figure 8. Event detection (STA/LTA used Allen's CF and k-mean CF) on the synthetic dataset containing 100 microseismic events. a) Detection threshold versus the number of triggers plot, showing that the smaller the threshold value, the higher the number of triggers. b) False positive rate (FPR) versus true positive rate (TPR) plot. c) Detection threshold versus TPR- FPR. The maximum value of which represents the Youden index and is considered the optimal detection threshold value ($\epsilon = 1.95$ for Allen's CF and $\epsilon = 3.05$ for k-mean CF) on this dataset.

158x54mm (300 x 300 DPI)

Figure 9. a) S/N of synthetic microseismic events. Higher S/N values are clipped to zoom in on the events with lower S/N. b) Moment magnitudes (M_w) that are uniformly distributed between -1 and -3. Detection curve (green) and 3C waveforms for c) optimal threshold value ($\epsilon = 1.95$) from Figure 8 for Allen's CF. Vertical arrows indicate the missing events. d) optimal threshold value (3.05) from Figure 8 for the k-mean CF. e) threshold value ($\epsilon = 4.0$) for Allen's CF. f) threshold value ($\epsilon = 4.0$) for the k-mean CF.

171x147mm (300 x 300 DPI)

Figure 10. Example of real data containing five microseismic events in which three have relatively low S/N waveforms.

141x110mm (300 x 300 DPI)

Figure 11. Example of event detection on real microseismic data. a) 3C waveform at receiver level 10 from the gather shown in Figure 10. Various bandpass filters (20-100Hz, 50-200Hz, 150-300Hz) were applied to data for extracting useful signal. b) STA/LTA responses computed using Allen's CF for the 3C input waveform in a). Horizontal lines represent the threshold value used to detect microseismic events (red = optimal threshold ($\epsilon = 1.95$) from synthetic examples, blue = threshold ($\epsilon = 3.05$). The rectangles are used to highlight areas of poor detections. The arrow shows that the optimal threshold value is too low for 150-300Hz filtered waveforms. c) STA/LTA responses computed using k-mean CF. Horizontal line indicates the optimal threshold value ($\epsilon = 3.05$) from synthetic examples in Figure 8. In this case, all five events are detected in all passbands and the number of false positives is lower in 150-300Hz filtered waveforms. The rectangles highlight the similar areas as in b) for comparisons.

105x248mm (300 x 300 DPI)

Figure 12. Plots of k versus peak STA/LTA response for microseismic events from the synthetic data example in Figure 7 at a) 1s, b) 2s, c) 3s, d) 4s, e) 6s, and f) 7s. Vertical lines indicate the signal's dominant period (dashed) and $1.5\tau_D$ (bold).

176x132mm (300 x 300 DPI)

DATA AND MATERIALS AVAILABILITY

Data associated with this research are confidential and cannot be released.