



Event detection using a fast matched filter algorithm - An efficient way to deal with big microseismic datasets

Hanh Bui, Mirko van der Baan

Dept. of Physics, CCIS, University of Alberta, T6G 2E1, Canada

Summary

Event detection is one of the time-consuming parts in microseismic processing. Different automated event detection algorithms have been proposed, such as the short-time average over the long-time average (STA/LTA), power spectral density, and subspace detection. However, these approaches are not convenient for big datasets. In this study, we introduce a fast matched filter (MF) algorithm, which can solve the efficiency challenge problem for these detectors. The proposed fast MF is built based on the fast normalized cross-correlation (NCC) computation technique. This method detects events in the data based on their similarity with the template events by comparing the NCC coefficients between the template events and the data with a specific user-defined threshold. The detection workflow is easy to follow with six main steps, namely data preconditioning, selecting template waveforms, multiplexing, fast NCC computation, extracting potential events, and quality control of the detection results. We have implemented the MF algorithm to a big microseismic dataset. The detection results obtained from the MF are compared with the results from the commonly used method, STA/LTA. The comparison shows that the MF algorithm is more efficient in event detection with fewer false alarms and higher detection probability within a shorter processing time than the STA/LTA, especially when dealing with big, noisy datasets.

Theory / Method / Workflow

Matched filter (MF) is a cross-correlation-based detection method. It is an effective method to search for a known signal (also known as a template event) in noisy data (Gibbons and Ringdal, 2006). The MF detects events based on their similarity with template events. It uses the template events extracted from the data and cross-correlates them with the continuous data to obtain an NCC coefficient matrix, C , representing the level of similarity between the template events and the data. Values of C will always be in between -1 and 1. A high absolute value of C means a high level of waveform similarity, and a low absolute value of C indicates little similarity between the two waveforms (Gibbons and Ringdal, 2006).

In this study, we build a fast MF algorithm, which is based on the fast NCC technique proposed by Lewis, 1995. The detection workflow includes 6 steps, namely (1) loading and filtering the input data, (2) extracting and selecting template events, (3) multiplexing data and templates, (4) fast NCC computation between template events and the continuous data using their multiplexed formats, (5) extracting the potential events when the NCC values exceed a trigger threshold, and (6) quality control (QC) of the detection results. In this workflow, we faster the detection process by using the STA/LTA algorithm with the recursive formula proposed by Withers et al., 1998 to extract the template events. The STA, LTA are then calculated as

$$STA_i = C_1 x_i + (1 - C_1) STA_{i-1} \quad (1)$$

$$LTA_i = C_2 x_i + (1 - C_2) LTA_{i-1} \quad (2)$$

in which $C_1 = 1/w_s$ and $C_2 = 1/w_l$; w_s , w_l are the short and long-time window lengths; x_i is the time series with the time index i . The recursive STA/LTA helps to avoid keeping long data vectors in the memory and reduce the computation time effectively (Withers et al., 1998). Then, the multiplexing technique is applied to turn the 3C data/templates into single continuous data

streams, so we only need to calculate the cross-correlation function on a single data stream and can reduce the complexity of the computation. Next, the fast NCC technique is applied to obtain the NCC coefficient matrix. The NCC coefficient between the template event T and the data $f(t)$ is given as

$$C(u) = \frac{\sum_t [f(t) - \bar{f}_u][T(t - u) - \bar{T}]}{\sqrt{\sum_t [f(t) - \bar{f}_u]^2 \sum_t [T(t - u) - \bar{T}]^2}} \quad (3)$$

where $C(u)$ is the NCC coefficient at each point u , T is the template event, \bar{T} is the mean of the template, \bar{f}_u is the mean of $f(t)$ in the region under the template. In Lewis's technique expressions in equation (3) can be efficiently computed with very few operations so we can speed up the detection process. After that, potential events are triggered and extracted when the NCC coefficient is higher than a user-defined threshold. Finally, quality control of the detection results is performed to remove undesirable events. This step is usually done by manual inspection and classification.

Results, Observations, Conclusions

To assess the detection performance of the proposed fast MF, we have implemented this algorithm on a big microseismic dataset (about 1.2 TB) and compared the detection results with the results obtained from the commonly used method, STA/LTA. The data are microseismicity emitted from 78 hydraulic fracturing (HF) treatment stages in 4 HF wells and are continuously recorded by sensors in both vertical and horizontal monitoring arrays. Figure 1 below shows the map view of the location of the treatment and monitoring wells. We run the STA/LTA with an STA window length being three times the dominant period of the event, an LTA window length being five times longer than the STA window, a trigger threshold of 2, and at least half number of receivers must observe the events. The proposed fast MF is implemented with a threshold of 0.2, and at least half of the number of receivers must see the events. Figures 2 and 3 below show the detection results obtained from both methods for each treatment stage. As we can see from these figures, the number of events obtained from the proposed MF and the STA/LTA in each HF stage is almost the same. After manual inspection and classification, we obtain a total of 21766 excellent events (those having both clear P- and S- phases) from the STA/LTA and a total of 19913 excellent events from the proposed fast MF. Thus, the proposed MF algorithm can detect almost the same number of events as the STA/LTA. However, this method requires less time for the detection process than the STA/LTA. The STA/LTA method is an incoherent energy detector which detects events without knowing information on the signals to be detected; thus, noise such as tube waves, electrical noise, and random noise can be incorrectly considered as potential events (those are false alarms/false triggers). With a threshold of 2, the STA/LTA helps to capture almost the number of true events (those having clear P- and/or S-phases) in the data; however, it has lots of false alarms in the detection results. Due to these false alarms, classifying the detection results in the STA/LTA method is time-consuming. In contrast, the proposed MF detects events based on their similarity with the template events. With the threshold of 0.2, the MF can detect almost the same number of excellent events while having fewer false triggers, which save time in the classification step. Furthermore, the combination of the recursive STA/LTA, multiplexing, and fast NCC computation techniques in the workflow fasters the detection performance of the proposed MF.

In summary, the proposed fast MF algorithm can work well with big microseismic datasets. The algorithm speeds up the detection process by applying the recursive STA/LTA combined with multiplexing and the fast NCC techniques. The workflow is easy to follow with a more superior

detection performance (less false triggers and high detection probability) than the commonly used method, STA/LTA. To perform the fast MF efficiently, we recommend using this algorithm for data generated from repetitive sources. If there is high variability in the waveforms, the MF can slow down the processing process as more templates need to be considered. The detection threshold can vary depending on the quality of the data. However, it should satisfy a trade-off between true events, false alarms, and missed events.

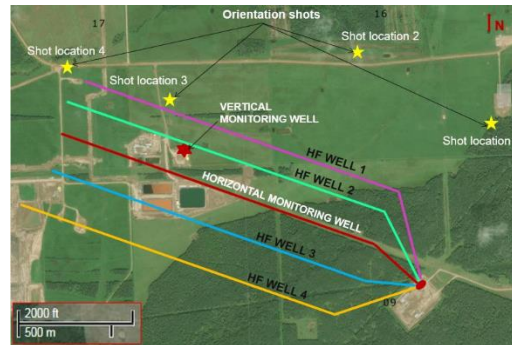
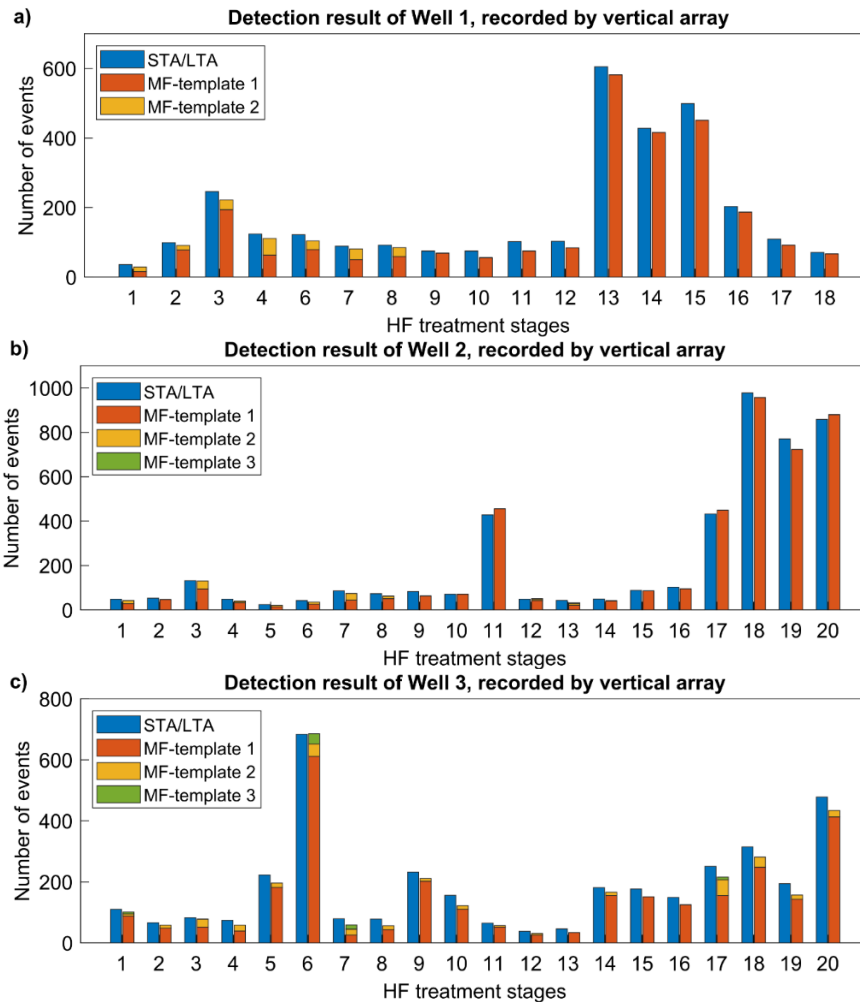


Figure 1 - Map view of the location of the HF treatment and monitoring wells.



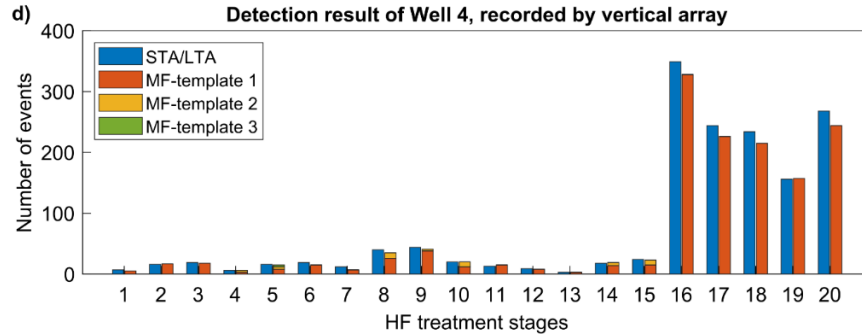


Figure 2 - The detection results of (a) Well 1, (b) Well 2, (c) Well 3, and (d) Well 4, recorded by the vertical monitoring array, obtained from the STA/LTA and the fast MF methods.

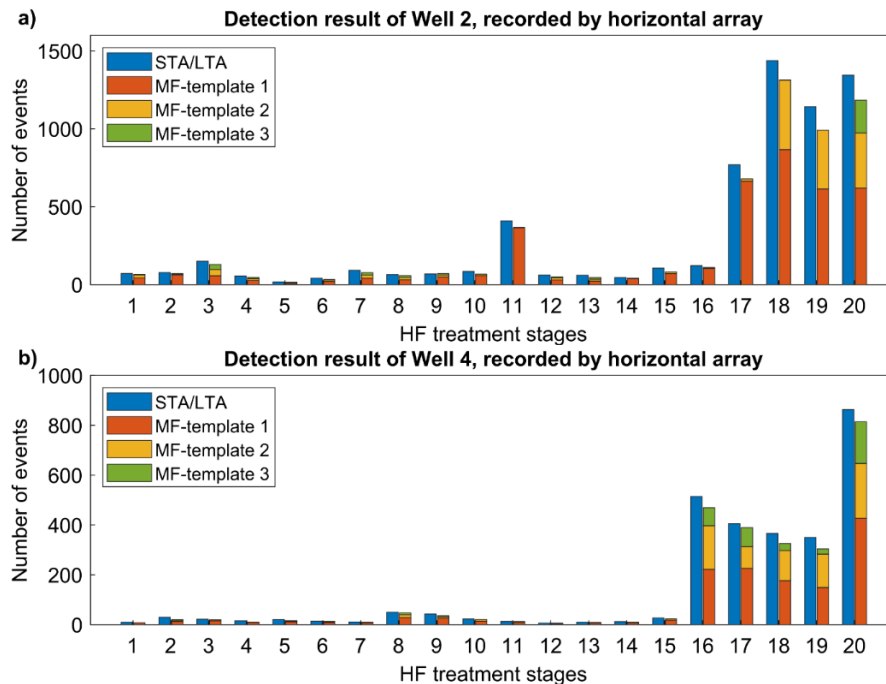


Figure 3 - The detection results of (a) Well 1, (b) Well 2, recorded by the horizontal monitoring array, obtained from the STA/LTA and the fast MF methods.

Acknowledgements

The authors would like to thank the sponsors of the Microseismic Industry Consortium for financial support and an anonymous company for permission to use and show the dataset.

References

- Gibbons, S. J. and F. Ringdal (2006). "The detection of low magnitude seismic events using array-based waveform correlation". In: *Geophysical Journal International* 165.1, pp. 149–166.
- Lewis, J. P. (1995). "Fast normalized cross-correlation". In: *Industrial Light & Magic*.
- 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". In: *Bulletin of the Seismological Society of America* 88.1, pp. 95–106.